Public Opinion Quarterly

Published for the American
Association for Public Opinion
Research by the University of
Chicago Press

TOM W. SMITH Liberal and Conservative Trends in the United States

JOEL E. BROOKS The Opinion-Policy Nexus in Germany

DONALD GRANBERG and SÖREN HOLMBERG The Berelson Paradox Reconsidered

ELIZABETH MARTIN, TERESA J. DEMAIO, and PAMELA C. CAMPANELLI Context Effects for Census Measures

PETER R. SCHROTT Electoral Consequences of "Winning" Debates

STANLEY PRESSER Can Context Changes Reduce Vote Overreporting?

STEVEN L. GORTMAKER and others Television Viewing and Mental Aptitude

THE POLLS Review: NBC's "R.A.C.E." *by Eleanor Singer*

Report: Finding Public Opinion Data *by Tom W. Smith and Frederick D. Weil*

BOOKS Reviews *by Seymour Sudman, Lawrence Bobo, Allan H. Spear, and Paul D. Reynolds*

CUMULATIVE INDEX TO VOLUMES 47–54, 1983–1990

Winter 1990

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR Advisory Committee

Eleanor Singer, Chair (1987–91)
Herbert Abelson (1988–92) David L. Sills (1986–90)
Gladys Engel Lang (1989–93)

Norman Bradburn and Andrew Kohut,
AAPOR Councillors-at-Large

Howard Schuman, *Editor*

Manuscript Editor John D. Vander Weg
Office Manager Carlotta Nosel
Assistant to the Editor Cheryl Rieger

The Polls
Reviews: Stanley Presser, Seymour Sudman
Reports: Tom W. Smith

EDITORIAL BOARD

Robert P. Abelson
Barbara A. Bailar
Sandra J. Ball-Rokeach
Larry M. Bartels
George F. Bishop
Lawrence D. Bobo
Jean M. Converse
Thomas D. Cook
Irving Crespi
James A. Davis
Stanley Feldman
Richard J. Fox
Norval D. Glenn
Robert M. Groves
Deborah R. Hensler
Shanto Iyengar
Jon A. Krosnick
Kurt Lang
Michael MacKuen
McKee J. McClendon
Jack M. McLeod
John Modell
John Mueller
Harold Nisselson
Helmut Norpoth
Steven J. Rosenstone
Norbert Schwartz
David O. Sears
Robert Y. Shapiro
Lee Sigelman

Public Opinion Quarterly (ISSN 0033-362X) Volume #54, 4 issues (1990)

© 1990 by the American Association for Public Opinion Research. Published quarterly: Spring, Summer, Autumn, and Winter, by The University of Chicago Press, 5720 S. Woodlawn Avenue, Chicago, Illinois 60637. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $40.00; individuals 1 year $20.00. WAPOR member subscription rate, U.S.A.: individuals 1 year $18.00. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $18.00. Other countries add $3.00 for each year's subscription to cover postage.

Japanese subscription agent: Kinokuniya Company Ltd.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, The University of Chicago Press, P.O. Box 37005, Chicao, IL 60637.

Mail orders and and direct subscription information to: Journals Fulfillment Department, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Institute for Social Research
P.O. Box 1248, Ann Arbor, MI 48016
Business, Subscriptions, Back Issues: The University of Chicago Press,
5801 South Ellis, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, New Jersey 08540
Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. ♾

PUBLIC OPINION QUARTERLY

VOLUME 54 / WINTER 1990 / NUMBER 4

Articles

Liberal and Conservative Trends in the United States since World War II 479
TOM W. SMITH

The Opinion-Policy Nexus in Germany 508
JOEL E. BROOKS

The Berelson Paradox Reconsidered: Intention-Behavior Changers in U.S. and Swedish Election Campaigns 530
DONALD GRANBERG and SÖREN HOLMBERG

Context Effects for Census Measures of Race and Hispanic Origin 551
ELIZABETH MARTIN, THERESA J. DEMAIO, and PAMELA C. CAMPANELLI

Electoral Consequences of "Winning" Televised Campaign Debates 567
PETER R. SCHROTT

Can Changes in Context Reduce Vote Overreporting in Surveys? 586
STANLEY PRESSER

The Impact of Television Viewing on Mental Aptitude and Achievement: A Longitudinal Study 594
STEVEN L. GORTMAKER, CHARLES A. SALTER, DEBORAH K. WALKER, and WILLIAM H. DIETZ

The Polls

Review: NBC's "R.A.C.E." 605
ELEANOR SINGER

Report: Finding Public Opinion Data: A Guide to the Sources 609
TOM W. SMITH and FREDERICK D. WEIL

Book Reviews

Robert M. Groves, *Survey Errors and Survey Costs* 627
SEYMOUR SUDMAN

Bob Altemeyer, *Enemies of Freedom: Understanding Right-Wing Authoritarianism* 629
LAWRENCE BOBO

Sidney Fine, *Violence in the Model City: The Cavanagh Administration, Race Relations, and the Detroit Riot of 1967* 633
ALLAN H. SPEAR

Bernard Barber, *Effective Social Science: Eight Cases in Economics, Political Science, and Sociology* 636
PAUL D. REYNOLDS

NOTICE TO CONTRIBUTORS 638

CUMULATIVE INDEX TO VOLUMES 47–54, 1983–1990 641

LIBERAL AND CONSERVATIVE TRENDS IN THE UNITED STATES SINCE WORLD WAR II

TOM W. SMITH

Abstract Analysis of 455 survey trends during the post–World War II period shows that America has generally moved in a liberal direction. The growth of liberalism has not been uniform across topics and time however. Trends dealing with equal rights and individualism had the most consistent liberal movement. Trends dealing with economic regulation and government power showed mixed change, while movement was mostly in the conservative direction on the topic of crime. Liberal growth was strongest during the 1960s and early 1970s. In the mid-1970s, many liberal trends slowed, with some stopping their advance and a few reversing direction. On average, liberal growth leveled off, but did not move in a conservative direction. In large part, this shift in social change represents a response to the events of the period, but it may also be a periodic alternation of the cycle of reform.

Studies of social change tend to be of two extreme types. At one end there are the general works on social change, especially sweeping studies covering the life and death of civilizations and discussing the process of modernization. These works are usually either broad, historical overviews or are theoretical (Applebaum, 1970; Boudon, 1983; Cairns, 1962; Cochran, 1972; Durant and Durant, 1968; Holton, 1981; Nisbet, 1969; Ogburn, 1966; Ryan, 1969; Sorokin, 1941; Spengler, 1926–28; Stinchcombe, 1978; Tominaga, 1982). At the other end are particularistic works that study one example or agent of change. These include monographs and case studies of particular actors, social movements, innovations, or time-series. This is particularly the case in stud-

TOM W. SMITH is Director of the General Social Survey at the National Opinion Research Center, University of Chicago. This research was done for the General Social Survey project directed by James A. Davis and Tom W. Smith. The project is funded by the National Science Foundation, Grant no. SES87-18467.

Public Opinion Quarterly 54:479–507 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-04/$2.50

ies of attitude change using surveys, where most time-series analysis either tracks a single question or a small cluster of closely related items (Smith and Crovitz, 1988).

While both the general and case studies clearly contribute to our understanding of social change, advancement has been hindered by the dearth of middle-range studies that empirically evaluate changes across a range of phenomena. This paper attempts to follow this middle course by examining changes in the preferences of Americans from World War II to the late 1980s on a wide range of liberal-conservative issues. Using survey time-series as the data blocks, the goal is to summarize trends, to look at the broad contour of social change, and to evaluate forces influencing liberal-conservative attitudes in contemporary America.

This research project emerged out of the interest of the National Data Program for the Social Sciences in monitoring and modeling social change. The first paper exploring liberal-conservative change was written in 1979 (Smith, 1982) and analyzed 111 time trends. Further work was done throughout the 1980s (Davis, 1980; Smith, 1984, 1985a, 1985b, 1985c, and 1988), expanding the data base and refining the analysis, until we presently have 455 time trends covering the period through 1987.

Liberalism/Conservatism

So what is this liberalism/conservatism that we seek to study? Unfortunately, liberalism is a hard concept to pin down and describe. Part of the problem is that liberalism is a chimera that has changed its emphasis and even some key tenets over time. Another impediment is the old problem of the blind men and the elephant. Liberalism is a big creature with many distinctive features and many observers describe single parts rather than whole. Finally, liberalism has evolved out of a long historical tradition and has many fathers. There is no definitive author or work that one can turn to as an authoritative pronouncement of what contemporary liberalism is. Because of such problems we will not try to offer a comprehensive definition or rigorous criteria for measuring liberalism. Instead we have (1) developed a list of liberal/conservative attributes, (2) examined how issues have been classified in contemporary politics since World War II, and (3) tested issues against measures of liberalism in surveys.

First, we examined many works on politics and political ideology for the prime attributes of contemporary liberalism (variously referred to as collectivist, interventionist, new, reform, or social liberalism to distinguish it from classical or laissez-faire liberalism).[1] The list we compiled is far from complete and ignores the problem that the attributes are not always complementary and are occasionally contradictory. With these limitations we found that, domestically, contemporary liberalism is (1) reformist, opting for change and generally opposed to the status quo; (2) democratic, favoring a full extension of electoral rights; (3) libertarian, supporting civil liberties such as free speech and the right to protest; (4) regulatory and interventionist, backing the management of business and the economy by the government; (5) centralist, using the federal government to set and enforce national standards and regulate state and local governments; (6) humanitarian, favoring a social welfare system for the care and protection of society in general and the lower class in particular; (7) egalitarian, advocating equal treatment for all and perhaps equal conditions for all; and (8) permissive, tolerating and often approving of nontraditional life styles and practices (e.g., homosexuality, nudity, and the use of drugs). In the foreign arena there is even less consensus on how applicable liberalism is and what liberal positions are. We found that, internationally, contemporary liberalism is (1) internationalist, supporting active U.S. involvement in the world; (2) multinational, backing the UN and other collective efforts; (3) nonmilitarist, preferring nonmilitary solutions to international disputes; and (4) prodétente, advocating good relations with Communist nations and not emphasizing anti-Communism as a cornerstone of foreign policy.

Next, we considered how issues had been defined in common usage in the post–World War II political scene. We considered the positions that various political parties, leaders, and groups had taken on these issues. In particular we considered how issues had been viewed by two archetypical liberal and conservative groups, respectively the Americans for Democratic Action (ADA) and the American Conservative Union (ACU) (Brock, 1962; Gillon, 1987; Libros, 1975). We determined whether the issues had been framed in liberal/conservative terms in the political arena and what positions, if any, liberal and conservative groups such as the ADA and ACU had taken on the issues.

Finally, we cross-tabulated all issues that appeared on the General Social Survey (GSS) with three indicators of liberal/conservative leanings of respondents: presidential vote in the 1972 and 1984 elections and self-placement on a seven-point liberal/conservative scale. For items that did not appear on the GSS we tried to find other criterion variables to verify (or challenge) our assessments. What was available varied from other attitude items that appeared on the GSS and had been themselves vetted to demographics such as party identification and race. Obviously the variables available to check our liberal/conservative judgments were sometimes less than ideal. In sum, if an issue related to one of the liberal principles listed above was considered to be a liberal/conservative issue by political actors and organizations during the post–World War II period, and was verified as related to liberalism by the survey data (or at least there was no negative evidence from survey cross-tabulations), we accepted the issue as tapping the liberal/conservative continuum.

---

[1] The works consulted included both historical and social science works on America's political tradition as well as specific studies on contemporary liberalism, including how citizens understand the terms liberal and conservative. Among the sources consulted are: Anagnoson, 1972; Chong, McClosky, and Zaller, 1983; Conover and Feldman, 1981; Gerber, 1975; Hamby, 1985; Hartz, 1962; Hero, 1969; Kelley, 1977; Lowi, 1969; Luttbeg and Gant, 1985; McClosky and Zaller, 1984; Maddox and Lilie, 1984; Mansfield, 1978; Pollock, 1983; Robinson and Fleishman, 1984; and Rotunda, 1986.

The list of survey items we compiled to measure liberal/conservative change are obviously neither definitive nor unproblematic. They represent an extensive range of political issues, but clearly miss or underrepresent many important issues. In addition to those items in the foreign policy arena that are generally less clearly and centrally related to liberalism than most domestic items, there are perhaps about two dozen items whose relation to liberalism is weak, tangential, or debatable. For example, we included support for Hawaiian statehood as a measure of liberalism since admission to the union was opposed by Southerners concerned about the territory's racial composition and also represented change rather than the status quo (Bell, 1984). Of course, Hawaiian statehood was hardly a central liberal/conservative issue. Overall, however, we believe that the issues we defined as liberal are both valid and reasonably representative indicators of liberal/conservative change since World War II.

The collection of time trends started with time-series in the General Social Survey and the American National Election Studies, but has been expanded to cover items from over a dozen different survey organizations. The time-series were identified through extensive searches in survey archives, especially the Roper Center for Public Opinion Research, University of Connecticut; the Inter-University Consortium for Social and Political Research, University of Michigan; the Louis Harris Data Center, University of North Carolina; and the National Opinion Research Center, University of Chicago.[2]

Three major steps were taken in evaluating each item for inclusion. First, did an item measure a liberal/conservative issue as discussed above? Second, was there a time-series for the item? Third, when could minor variations in wording be ignored so that a single time-series could be constructed? Obviously we preferred time-series made up of only absolutely equivalent wordings. Fortunately in a majority of cases we were able to use only verbatim items. Often, however, a potential series included questions with various minor differences in wording. When the questions were judged to be equivalents, the minor differences in wording were accepted and one time-series was created.[3] All variations in wording are documented in the data base that holds these time-series. Figure 1 shows an example of a question in the data base that has several variant wordings.

---

[2] For a list of the chief archival guides and sources see Davis and Smith, 1989: appendix N.

The data base of 455 time-series has various pluses and minuses. On the one hand it is by far the most extensive collection of survey trend data ever assembled and analyzed together. There are an average of 8.1 readings per series for a total of 3,707 data points and the time-series cover an average of 14.8 years. Yet this is far from total coverage. Over the 43 years from 1945 through 1987, the 455 time-series could have a total of 19,565 data points (43 × 455) and we have points for only 3,707 or 18.9%. The shortfall comes largely because most time-series cover only a part of the time period. Looking at periods of about 15 years, we see that 13.2% of the time-series start in 1936–45, 23.3% in 1946–59, 47.3% in 1960–74, and 16.3% in 1975 or later. Considering both the starting and stopping points in the time-series, we find only 4.0% fit the ideal condition of starting by 1945 and ending after 1974, a plurality of trends (39.6%) started in 1960–74 and ended after 1974, and 16.3% both started and ended since 1974.

The 455 time-series consists of 419 attitudes on which respondents express their personal preferences or beliefs, 13 nonaffective judgments in which respondents report on what they believe trends are or how society is (e.g., "Do you think drunkeness is increasing or decreasing in this community?" or "Do you think the day will ever come in the South when whites and Negroes will be going to the same schools, eating in the same restaurants, and generally sharing the same public accommodations?"), and 23 measures of personal or household behavior (e.g., "Have you seen an X-rated movie in the last year?" or "Do you (or does your husband/wife) go hunting?").

The 455 time-series also cover a wide range of issues. We organized the items into 17 categories that represented major issue areas com-

---

[3] First, based on knowledge of the literature on survey methods, a judgment was made whether the variation in wording was sufficient to alter the response distribution. Second, each time-series with variant wordings was inspected for outliers or blips that might be due to wording. According to these two criteria, certain variant wordings were included in a time-series. For treatments of special problems with using survey data to study change over time see Bahr, Caplow, and Chadwick, 1983; Glenn, 1975; Hyman, 1972; Kulka, 1982; and Martin, 1983.

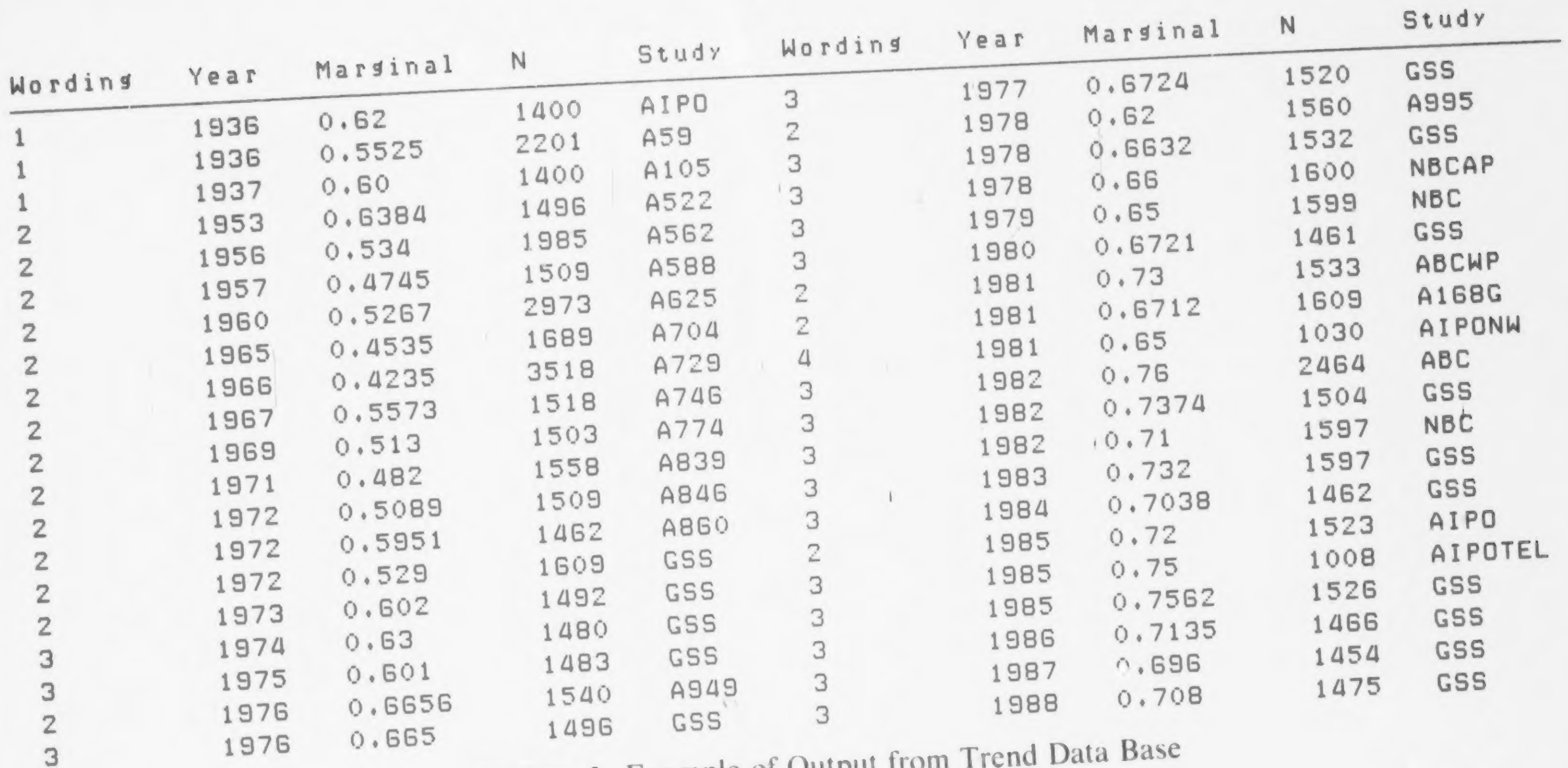
CAPPUN

1 Are you in favor of the death penalty for murder.
2 Are you in favor of the death penalty for persons convicted of murder.
3 Do you favor or oppose the death penalty for persons convicted of murder.
4 Are you in favor of or opposed to the death penalty for persons convicted of murder.

Cut: Proportion equals Favor versus Oppose and DK

TEST	ESTIMATE	CHI SQ	PROB
Constant	0.625 0.0038	2642.2	<.001
Linear Trend			
Weighted Regression	Y = −7.95 + 0.0044 (X)		
R Squared	0.3303		
Improvement		944.4	<.001
Fit		1697.8	<.001

Model: Significant Linear Component



```
Wording    Year    Marginal    N       Study    Wording    Year    Marginal    N       Study
1          1936    0.62        1400    AIPO     3          1977    0.6724      1520    GSS
1          1936    0.5525      2201    A59      2          1978    0.62        1560    A995
1          1937    0.60        1400    A105     3          1978    0.6632      1532    GSS
2          1953    0.6384      1496    A522     3          1978    0.66        1600    NBCAP
2          1956    0.534       1985    A562     3          1979    0.65        1599    NBC
2          1957    0.4745      1509    A588     3          1980    0.6721      1461    GSS
2          1960    0.5267      2973    A625     2          1981    0.73        1533    ABCWP
2          1965    0.4535      1689    A704     2          1981    0.6712      1609    A168G
2          1966    0.4235      3518    A729     4          1981    0.65        1030    AIPONW
2          1967    0.5573      1518    A746     3          1982    0.76        2464    ABC
2          1969    0.513       1503    A774     3          1982    0.7374      1504    GSS
2          1971    0.482       1558    A839     3          1982    0.71        1597    NBC
2          1972    0.5089      1509    A846     3          1983    0.732       1597    GSS
2          1972    0.5951      1462    A860     3          1984    0.7038      1462    GSS
2          1972    0.529       1609    GSS      2          1985    0.72        1523    AIPO
2          1973    0.602       1492    GSS      3          1985    0.75        1008    AIPOTEL
3          1974    0.63        1480    GSS      3          1985    0.7562      1526    GSS
3          1975    0.601       1483    GSS      3          1986    0.7135      1466    GSS
2          1976    0.6656      1540    A949     3          1987    0.696       1454    GSS
3          1976    0.665       1496    GSS      3          1988    0.708       1475    GSS
```

Figure 1. Example of Output from Trend Data Base

monly discussed and defined in the media and politics. These topics generally reflect the face purpose of the question (e.g., it is a race relation question or an item about the military), rather than any underlying dimensions or principles that the question might refer to (more of this alternative approach later). Classifying items into such topics was, of course, sometimes difficult because a question might cover two or more topics (e.g., a question on the poll tax dealt both with racial equality and democratization and an item on military spending obviously fit into both the military and spending/taxation categories), but most of the items clearly fit into the topics they are classified under. Overall for the 17 categories we employed, the 455 time-series were distributed as follows:

Abortion	2.0%
Civil Liberties	9.7
Crime	5.3
Economic Regulation	7.3
Family	1.3
Feminism	8.4
Foreign Affairs	3.7
Labor	6.6
Lifestyles	2.0
Military Recruitment	2.2
Miscellaneous	3.5
Politics	5.1
Race and Ethnicity	14.9
Religion	5.7
Sex	6.2
Social Welfare	5.3
Spending and Taxes	11.0
N	455

To categorize each trend, a series of models are fitted to the data points. First, we try the constant model that all data points are simple random variations around a stable proportion. An estimate of the pooled (or average) proportion is made and we test to see if the observed data points vary significantly from it. If this model is rejected, the linear model that all data points are random variations around a linear trend is tried. Four outcomes are possible: (1) constant, (2) significant linear component, (3) significant linear trend, and (4) nonconstant, nonlinear. The constant model is accepted when there is no significant variation around the constant or pooled proportion. The significant linear component model is accepted when (*a*) the constant model is rejected, and (*b*) the linear model is rejected, but (*c*) the linear fit is significantly better than the constant fit. The significant linear trend model is accepted when (*a*) the constant model is rejected and (*b*) there is no significant variation from the linear model. The nonconstant, nonlinear model is accepted when (*a*) the constant model is rejected, (*b*) the linear model is rejected, and (*c*) the improvement between the constant and linear models is not significant (Taylor, 1980). Time-series that have either significant linear components or significant linear trends and are in the liberal direction are classified as liberal trends and, similarly, those in the conservative direction are classified as conservative trends.[4]

Across the 455 trends we find that 9.7% are constant, 11.2% are nonconstant, nonlinear (NCNL), 54.7% are significant linear components, and 24.4% are significant linear trends. When we classify the trends as either liberal or conservative we find that 20.9% have no direction (constant or NCNL), 54.9% are in the liberal direction, and 24.2% are in the conservative direction. Liberal trends thus exceed conservative trends by over two-to-one (2.27:1). Taking the percentage with a liberal trend minus the percentage with a conservative trend as a summary measure of the direction of liberal/conservative change, we find a +30.7% liberal/conservative score (with +100.0% meaning that all series showed trends in the liberal direction and −100.0% that all series had conservative trends).

Given the nonrepresentative nature of the time-series covered with their mixture of different topics and time periods, we might naturally wonder whether such a summary characterization is robust. We therefore recalculated the distribution of trends using several different adjustment procedures. The first adjustment procedure substituted a design effect of 2 for the SRS assumptions used in the initial calculations. This adjustment mainly increases the number of trends fitting the constant model. The second adjustment converted all constant models into linear trends by accepting the best linear fit for each series even if the constant model had initially been accepted. This allows us to see if essentially constant series show any directional leanings. The third series employed the opposite strategy by taking weak significant linear components and counting them as showing no trends. This discounts weak trends. The fourth weighted each trend by its interval (i.e., the number of years between first and last data point). This obviously gives more weight to long trends than to short time-series. The fifth adjustment gave equal weight to each of the topical areas (discussed below). This balancing of different topics is not any more representative than the self-weighting, but does allow us to see the impact of a notable shift in the relative size of topics. The sixth adjustment excluded from the analysis questions on foreign affairs and military topics. These exclusions remove the topics that are nondomestic and which, as a group, are hardest to fit consistently into a liberal/conservative framework over time.

---

[4] A table is available from the author that summarizes the content and direction of the 455 time-series. In addition, full documentation of the wordings, data points, sample sizes, etc., of all trends is available for the cost of duplication and shipping.

None of these adjustments, nor certain variations on them discussed below, had a major impact on the overall liberal/conservative score (table 1). The elimination of weak significant linear components lowers the liberal/conservative score the most, but its impact on the relative balance of liberal/conservative trends is somewhat less than first appears, since the ratio of liberal to conservative trends only falls from 2.27:1 to 2.11:1. The duration adjustment moves the liberal/conservative score about an equal margin in the opposite direction. The liberal gain reflects the fact that time-series showing liberal trends are longer than their conservative counterparts. Since the unadjusted figures are in the middle of the range of adjusted figures and none of the figures have any clear superiority over others, we have used the unadjusted figure elsewhere in the paper (except where noted).

A second summary measure of the direction of change is the net slope. This takes the best linear slope for each series and assigns a positive slope to liberal trends and a negative slope to conservative trends. Overall the net slope is 0.0040 per annum (a yearly liberal gain of 0.4 percentage points). This liberal edge comes entirely from the preponderance of liberal trends over conservative trends since the average slope for liberal and conservative trends was virtually identical (respectively + 1.32 and + 1.31 percentage points per annum).[5] We subjected this net slope estimate to various adjustment procedures similar to those applied to the liberal/conservative score and, as above, found little variation in the overall averages. We did find, however, that short-term trends sometimes based on only a few data points did exercise an inordinate influence in the unadjusted figures and in some of the subsequent analysis have weighted by the time-series interval to compensate for this.

While the liberal and conservative trends are similar in magnitude, the liberal trends have somewhat better linear fits than the conservative trends. Liberal trends have an average R^2 with time of .58 and the percent to total chi-square accounted for by the linear model is 73% compared to an R^2 of .47 and a percent of linear chi-square of 56% for conservative trends. This suggests that the liberal trends have more

---

[5] Since both constant trends and NCNL trends showed very marginal conservative leanings (respectively − .04% and − .09%), liberal and conservative trends are also almost equal in magnitude over all time-series.

Table 1. Liberal/Conservative Scores

Adjustment	Score
Weak significant linear components eliminated	+24.6
Equal weight to each topic	+28.8
Adjusted for design effects	+30.5
Unadjusted	+30.7
Foreign and military eliminated	+31.3
Constants counted as trends	+31.6
Weighted by duration	+37.3

strength than the conservative movements. The greater regularity of the liberal trends may result from the fact that they are more frequently driven by a cohort-education model that promotes steady liberal gain for many items (Davis, 1980; Pollack, 1983). In addition, the liberal trends cover longer intervals (17.1 years) than do conservative trends (13.0 years).

In terms of the "big picture," the view in post–World War II America has mainly been to the left. Liberal and conservative trends have been about equal in magnitude, but liberal trends have outnumbered conservative trends by over two to one. This, in turn, meant on average across all time-series there was a net liberal gain of about 0.4 percentage points per annum. These overall averages mask as much as they reveal however, because change has not been uniform across topics or across time.

Trends by Topics

Table 2 shows the trends for each of our 17 topics. The liberal/conservative scores are reasonably similar in both rank order and magnitude for all adjustments. There is, however, considerable range in the net direction of trends across topics, ranging from abortion, race/ethnicity, and feminism with scores of + 70.0 down to scores of − 20.0 for military recruitment, miscellaneous, and crime. In fact, the range is so heterogeneous by topic that no topic even falls close to the over-all score of + 30.7.

Table 3 shows the net slope for each topic, both unadjusted and weighted for length of time-series. Weighting by length of time-series does effect the relative ranking of topics and their magnitudes, and overall the weighted figures give a better description of the general

Table 2. Trends by Topical Areas

	Liberal/ Conservative Score	Liberal Direction	Trend		Conservative Direction	Total
			Constant	NCNL		
Abortion	77.8	77.8%	11.1	11.1	0.0	100.0%
Race/ethnicity	72.0	83.8	1.5	2.9	11.8	100.0
Feminism	71.1	73.7	18.4	5.3	2.6	100.0
Family	66.7	66.7	0.0	33.3	0.0	100.0
Religion	53.9	65.4	7.7	15.4	11.5	100.0
Foreign affairs	47.1	64.7	0.0	17.6	17.6	100.0
Lifestyles	44.5	55.6	0.0	33.3	11.1	100.0
Sex	42.8	57.1	14.3	14.3	14.3	100.0
Civil liberties	40.9	61.4	4.5	13.6	20.5	100.0
Economic regulation	10.2	57.6	3.0	0.0	39.4	100.0
Labor	10.0	43.3	13.3	10.0	33.3	100.0
Political	8.7	52.2	0.0	4.3	43.5	100.0
Spend/tax	−4.0	26.0	26.0	18.0	30.0	100.0
Social welfare	−8.4	33.3	8.3	16.7	41.7	100.0
Military recruitment	−20.0	30.0	10.0	10.0	50.0	100.0
Miscellaneous	−20.7	25.0	18.8	12.5	43.8	100.0
Crime	−20.8	25.0	12.5	16.7	45.8	100.0
All	30.7	54.9	9.7	11.2	24.2	100.0

NOTE: Liberal/Conservative Score = % liberal direction − % conservative direction.

Table 3. Net Slope by Topical Area (percentage points change per annum)

Topic	Unadjusted	Weighted by Interval
Race/ethnicity	1.16%	1.03%
Sex	0.91	0.62
Feminism	0.88	0.95
Abortion	0.72	0.84
Foreign affairs	0.62	0.38
Family	0.50	0.58
Civil liberties	0.46	0.75
Economic regulation	0.39	0.20
Military recruitment	0.26	0.22
Religion	0.22	0.28
Miscellaneous	0.08	0.03
Lifestyles	0.04	0.17
Social welfare	−0.00	0.03
Politics	−0.04	0.24
Spending/taxes	−0.06	0.13
Crime	−0.09	−0.15
Labor	−0.42	0.15
All	+0.40	+0.47

trends.[6] As in the case of the liberal/conservative scores in table 2, there is considerable range across topics in the magnitude of trends, ranging from racial/ethnic trends averaging over + 1.0 percentage points to crime at − 0.15 percentage points.

Table 4 compares the ranking of topics according to the liberal/conservative scores and net slopes. Overall the level of agreement is high. Race/ethnicity, feminism, and abortion are the most liberal on both lists, while social welfare, spending and taxes, miscellaneous, and crime are the most conservative on both lists. In the middle there is more disagreement between the two indicators, with sex and lifestyles the most inconsistent followed by religion and labor. The basic reason for the difference in rank for these topics is that (*a*) they contain a mixture of liberal and conservative trends and (*b*) the liberal and conservative trends differ in magnitude. The liberal/conservative score gives equal weight to each trend, while the net slope weights by the

---

[6] This assessment is based largely on the fact that a number of large changes were based on short time-series. These often have a large impact on topical averages, especially for those topics that contain relatively few time-series.

Table 4. Comparison of Rank Order of Topics by Liberal/Conservative Score and Net Slope

Liberal/Conservative	Net Slope
Abortion	Race/ethnicity
Feminism	Feminism
Race/ethnicity	Abortion
Religion	Civil liberties
Civil liberties	Sex
Lifestyles	Family
Family	Foreign affairs
Foreign affairs	Religion
Labor	Politics
Politics	Military recruitment
Sex	Economic regulation
Military recruitment	Lifestyles
Economic regulation	Labor
Social welfare	Spending/taxes
Miscellaneous	Social welfare
Spending/taxes	Miscellaneous
Crime	Crime

size of the slope. For example, lifestyles has only six trends and five small liberal trends are offset by one large conservative trend, so its rank using net slopes is six positions lower than on the liberal/conservative score ranking. Similarly on sex morality, the liberal trends are larger than the conservative trends, so the topic ranks as more liberal on the net slope list than on the liberal/conservative scale.

Examination of the rank order and magnitudes in tables 3–5 allows us to draw some tentative conclusions about factors influencing liberal/conservative change. First, the topics showing the strongest liberal change (race/ethnicity, feminism, and abortion) all primarily deal with increased support for freedom of individual choice and/or equal treatment/rights for individuals (regardless of such fixed attributes as gender and race). The most conservative topic, crime (Smith, 1976 and 1980), involves increased support for social control and in many ways might be seen as a counterbalance to the liberal trend toward more individual freedom. Generally in the middle one finds topics dealing with government versus private regulation over material matters (economic regulation, labor, spending and taxes, and social welfare).

Several of these middle-ranked topics show bimodal distributions (i.e., they show a notable proportion of trends in both the conservative and the liberal directions). Such a pattern might result from (*a*) different movement among topical subgroups, (*b*) different movement over different periods, (*c*) some combination of these two causes, or (*d*) some other factor, such as ambivalence or cross-cutting pressures. The political topic clearly follows the subgroup model with items about political rights and democratization showing almost uniform liberal movement, while items about government power and size show a conservative shift. This intratopical split reflects the intertopical difference we noted above under which extensions of individual freedom and equal rights grew while support for economic regulation and government control items lagged behind. Labor's bimodal trends are generally a result of period effects. Trends in the late 1940s and in the late 1970s to early 1980s were typically antilabor, while those in between were prolabor (more on period effects below). Economic regulation and social welfare show more complex patterns. Support for economic regulation increased for profit control, but fell for government ownership. On price control measures, it was high during periods of peak inflation and lower when inflation subsided. For social welfare there seems to be some real ambivalence, with different measures of support for education, health care, welfare, unemployment, and retirement moving in opposite directions. This mixed pattern is equally true if the spending items having to do with social welfare are considered as part of the social welfare topic. It suggests that Americans during this period were attracted by the benefits of entitlements, the appeal of humanitarianism, and the idea of human capital investment, but leery of both costs and government control.

Trends on Alternative Dimensions

To further explore what factors influenced or shaped liberal/conservative change, we evaluated items along three additional dimensions (i.e., besides liberalism/conservativism): (*a*) government involvement, (*b*) equal treatment/rights, and (*c*) individual choice. Each of the 455 time-series were independently rated as being related to each of these dimensions. While all 455 items were related to the general liberal/conservative dimension, only a subset of items were related to these three dimensions: government intervention (282), equality (115), and choice (191). Although there are some definite relationships between some of these dimensions and both themselves and the liberal/conservative dimension, each was separately created without reference to the other classifications.

Turning to government intervention first, we find that 14.5% of the applicable time-series showed no trend (constant or NCNL), 30.0% moved toward more state involvement, and 44.5% toward less state activity for an antistate balance of + 14.5 percentage points. Overall the government intervention dimension had little relationship to liberalism/conservatism. Conservative trends were in the prostate direction 37.5% of the time, while liberal trends moved toward more state activity in 42.3% of the cases. The antistate trends concerned topics and items in which more government would restrict individual rights and freedom of choice—abortion, civil liberties, feminism, and religion (requiring school prayer). Trends favoring state intervention were either to protect equal rights, especially of "powerless" groups (minorities and the working class) or to reestablish social control (crime and sex education in the schools). Generally, support for material entitlements and benefits (social welfare and spending/taxes) were not in the progovernment direction, although often the cross-sectional marginals indicated progovernment majorities.[7]

Movement among items related to the equal treatment dimension, which we discussed above both to explain the pattern of intertopical differences and differences on the role of government dimension, has been overwhelmingly toward favoring more equal treatment. No direction (constant/NCNL) emerged for 15.7% of the time-series, 71.3% were toward equality, and 13.0% moved away from equal treatment, for a more equality edge of + 58.3. Unlike government involvement, which was essentially unrelated to the liberal/conservative dimension, pro–equal rights trends were all classified as liberal trends. Thus, this dimension is merely a subset of liberal/conservatism and not an unrelated dimension. Items making up the set of equal rights items are overwhelmingly from three topics—race/ethnicity (48), civil liberties (29), and feminism (27). The few items that buck the general proequality trend are some civil liberty items from the 1940s and 1950s (7) and a few race/minority items (6) that called for special treatment (e.g., affirmative action) or spending for minority groups. Such group benefits racial items are of course seen as racist, anti-individual, and illiberal by some critics.

Trends involving items tapping freedom of choice show a general movement toward more support for individual choice, although not as strongly as the equal rights movement. No direction was shown by 23.6% of the time-series, 57.6% moved toward more freedom, and 18.8% were for less freedom. The net index score was 48.8 percentage points in the freedom direction. This dimension is more closely related to the liberal/conservative dimension than government intervention is, but is not a mere subset as equal rights is. Among conservative trends, 52.3% were in the individual choice direction, while 85.3% of liberal trends moved toward more freedom of choice. This dimension drew items from more topics than the equality dimension, but had very few items from foreign affairs (0), spendings/taxes (0), religion (1), politics (1), social welfare (4), or crime (4). Freedom of choice was generally expanded for ascripted groups such racial minorities and women (including the right to an abortion) and in the areas of sexual morality and for some lifestyles and miscellaneous issues (e.g., euthanasia, suicide, drugs, and prohibition). Movement against free choice was centered in three areas: civil liberties in the 1940s and 1950s, support for universal military training and the military draft, especially in the 1950s, and economic regulation (i.e., against free enterprise).

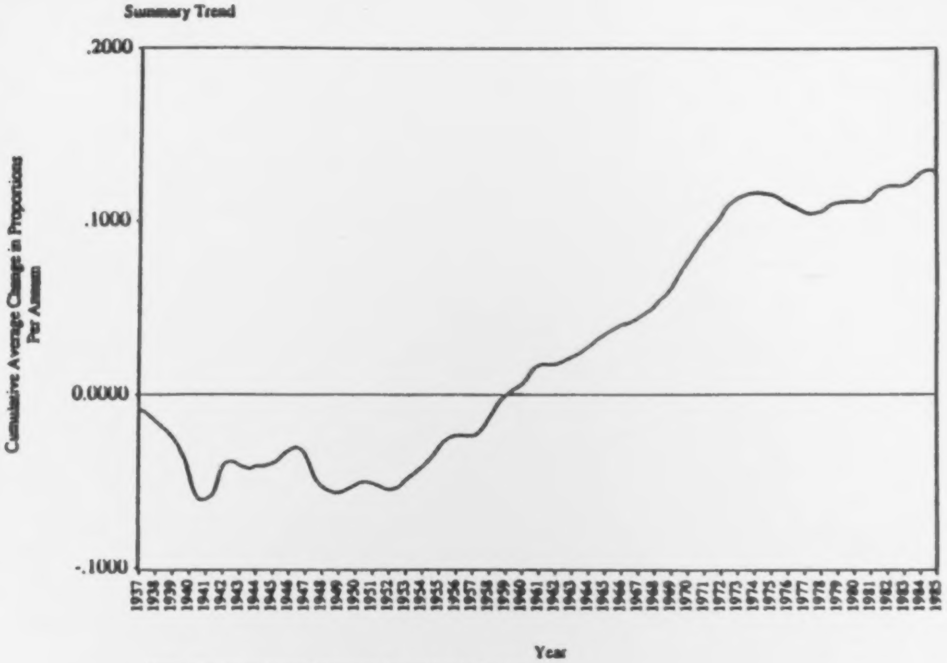
---

[7] This was especially true of spending items in the 1970s and early 1980s.

In brief, liberalism has prevailed over conservatism since World War II largely because of a widespread increase in public support for equal rights and individual choice (outside of the economic realm). Support for government control over material well-being via economic regulation and an expansion of social welfare has been less consistent in direction and lower in magnitude. The one area to clearly show conservative movement has been social control and crime. Other explanations and elaborations of the general pattern of change involve differences by subperiods.

Trends by Periods

We would like to be able to describe year-to-year changes in the annual rates of liberal/conservative change, but this is extremely difficult because of the general limitations of the individual time-series (e.g., mainly the differences in the starting and ending points of various series and internal gaps in each individual series) and, in particular, because of the shortfall of data points in the earlier years. We have made one preliminary stab at showing what the overall rate of change was by period in figure 2. By assuming that all change between data points was even (e.g., that a 10 percentage point shift over a five-year interval represented five annual gains of 2 percentage points), that an overall trend could be computed by taking the signed average of all estimated trends covering a given year, and that the mixture of trends represented from year-to-year does not significantly bias the annual averages, the average liberal trend emerges as depicted in figure 2. It suggests a fairly steady level of liberal gain from the early 1950s until about 1973–74 and then a leveling off of liberal movement through the mid-1980s. We refer to this leveling off in the mid-1970s as the "liberal plateau" to emphasize that, on balance, neither liberal nor conservative change predominated and that this flattening occurred after a long



Note: Maximum interval between data points is five years.

Figure 2. The Liberal Plateau in America, 1937–1985

liberal climb. (The trend line prior to 1945 should be entirely ignored because of the inadequacies in the number of time-series covering those years.) This overall trend should be considered only a first-cut approximation and steps are being taken to test its robustness, strengthen the data base, and refine the estimate.[8] The mid-1970s liberal plateau description is buttressed, however, by numerous individual examples, particularly in the areas of civil liberties and abortion (fig. 3), and there are sufficient data to explore this hypothesis more rigorously.

Among our 455 time-series there are 176 that have trends both before and after 1974 (i.e., two or more data points in both periods; including 13 time-series with only a total of three points with an observation in 1974 counting for both the pre- and postperiods). These 176 time-series have an average duration of 21.2 years and 13.0 data points. On one hand this is a relatively large, diverse, and rich set of time-series. However, it represents less than half of our initial time-series and covers a different mixture of topics. Looking over all years, these 176 time-series do show about the same mixture of no direction and liberal/

---

[8] For example, we are working on a procedure for merging closely related time-series via calibration in order to construct longer and denser time-series as well as to test the robustness of our present estimate.

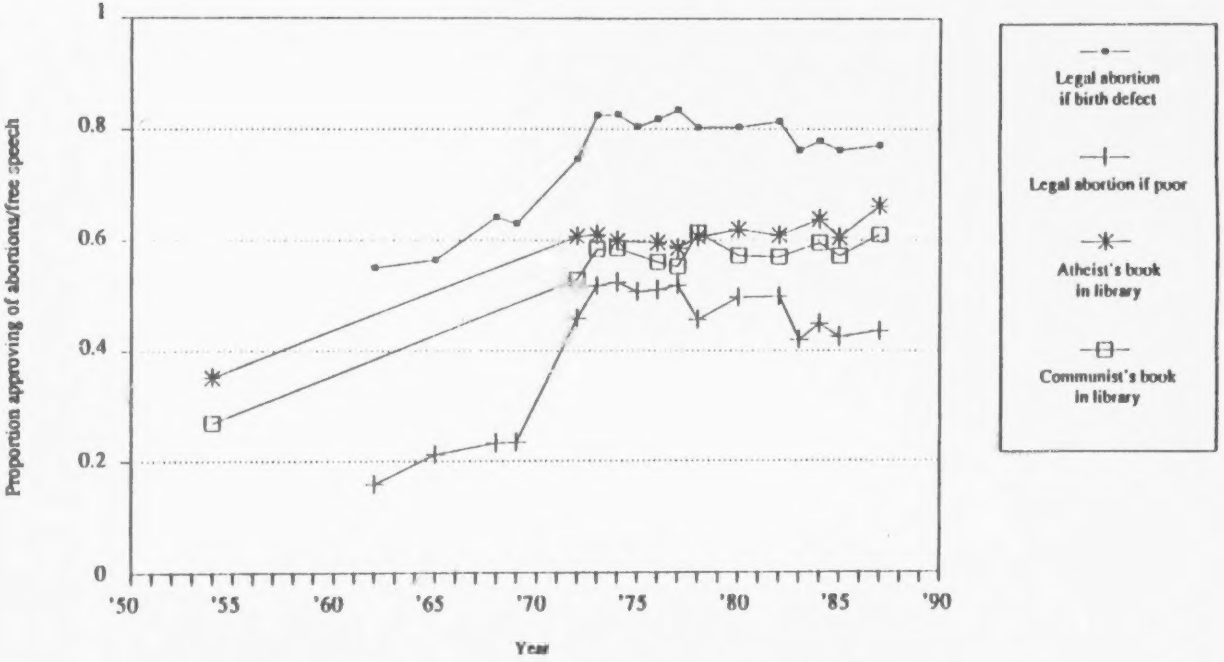


Figure 3. Selected Trends Demonstrating the Liberal Plateau

conservative movement as did the 455 items. Taking the unadjusted series, 2.3% were constant, 11.9% were NCNL, 26.1% moved in a conservative direction, and 59.7% in a liberal direction, for a liberal/conservative score of + 33.6. We applied a topical weight to the data so that each topic would represent the same share among the 176 series as it had among the 455 and this raised the liberal/conservative score slightly to + 36.5. For the net slope the unadjusted average was + 0.32 percentage points per annum and + 0.36 when weighted by topic. This puts the liberal/conservative score slightly above the overall average, while the net slope is a bit lower. In general, we find that the 176 trends are both reasonably representative of the 455 time-series and an adequate set of items for considering how liberal/conservative attitudes have varied over time.

Table 5 shows that there was a decided shift among trends before and after 1974. Prior to 1974, the trends were overwhelmingly liberal (+ 43.2), while since 1974 there has been almost an even balance between liberal and conservative trends (+ 4.0). This same pattern emerges when we look at net slopes. Until 1974 the net slope was 0.64 percentage points per annum, while since 1974 it has been − 0.04 percentage points.

Looking at the individual switches in the time-series, we find that 46.6% moved in a conservative direction (i.e., from a liberal trend to no direction, no direction to a conservative trend, or from a liberal to a conservative trend), 31.1% showed no change in direction, and 22.1% shifted in a liberal direction (i.e., a conservative trend to no direction, no direction to a liberal trend, or from a conservative to a liberal trend). Conservative shifts were most common on the abortion topic, which had a 100% turn around, followed by economic regulation, foreign affairs, and religion, which showed conservative shifts in over one-half of all possible cases. Shifts were least common among feminism items, with only 16.7% turning toward the conservative, followed by race/ethnicity and sex. Given that abortion and feminist attitudes are closely related issues and that these attitudes show moderately strong associations in cross-sectional surveys, it is striking that their trends, while consistent in direction prior to 1974, have diverged in recent years. This is not only an important finding in its own right, but cautions that related series do not always exhibit the same patterns of change.

Among the items showing ideological shifts, the most interesting are complete reversals, trends that were moving in one direction, but flipped sign to move in the opposite direction. Thirty-six liberal trends in the pre-1974 period changed to conservative trends since 1974, and 10 initial conservative trends became liberal. These 46 trends represent reversals for fully 36.5% of the possible cases (i.e., of the 126 time-

Table 5. Trends Before and After 1974

Model	To 1974	Since 1974
Constant	18.8%	22.2%
NCNL	9.7	9.1
Conservative trend	14.2	32.4
Liberal trend	57.4	36.4
Liberal/conservative score	+43.2	+4.0

series that had directional trends in the pre-1974 period). Liberal reversals were concentrated in a few areas and almost totally absent in others. They occurred on all six abortion items and were also concentrated among social control items (crime, drug control, and family preservation by limiting divorce), and government control and state welfare items. In contrast, only a single trend on race/ethnicity and feminism reversed.

Conservative reversals are less reliable, fairly scattered, and more particularistic. Many of the conservative reversals involve time-series with few data points and/or weak trends pre-1974, post-1974, or both. Those trends that are more substantial are not concentrated in any topic and do not seem to represent any common themes. Instead they seem to involve reactions to specific events and developments, such as decrease in fear of crime after a peak in the crime rate in the early 1980s, and a decline in concern over the power of government in the Reagan years.

In brief, the balance of liberal/conservative change since World War II has not been uniform over time. Many liberal trends leveled off and some even reversed direction in the 1970s. There are two notable features of this liberal plateau period. First, that overall society did not reverse from liberalism to conservativism, but the liberal gains of the post–World War II period did level off. While some liberal (and some conservative) trends did reverse direction, the general shift was from liberal advance to a liberal holding-pattern (Astin, Green, and Korn, 1987; Burnham, 1980; Chafetz and Ebaugh, 1983; Hastings and Hoge, 1981; Hoge, Luna, and Miller, 1981; Ladd, 1983; Moberg and Hoge, 1986; Thornton, 1985). Second, this leveling off did not happen after Reagan's victory in 1980 or even during 1978 with Republican election gains and the Proposition 13 tax revolt, but earlier in the 1970s, around 1974. Of course some trends swung to the conservative side earlier than that (e.g., support for capital punishment began increasing in the mid-1960s) or later (support for the legalization of marijuana leveling off in 1978–80 and dropping thereafter) and many other trends never did show significant declines (much less reversals). Still a number of trends showed points of inflection in 1974 or centered around that year.

Causes of Liberal/Conservative Change

In our earlier writings (Smith, 1982), we proposed a series of general to specific explanations for what has been called the general liberal hypothesis. To summarize these briefly, we argued that two basic long-term forces have moved America in a liberal direction: modernization and liberal idealism. Modernization promoted liberalism through the growth of rationalization, innovation, centralization, statism, and prosperity. Liberal idealism reflects the fact that, in our historical tradition and government roots, America is a liberal nation. The ideals symbolized in our political culture and enshrined in patriotic emblems from the Pledge of Allegiance to the Gettysburg Address are liberal ideals. While practice has often fallen far short of these ideals, the ideals have remained cherished and have continually exercised a pressure (sometimes latent and sometimes active) toward liberal change (Abbott, 1981; Hartz, 1962; McElvaine, 1987).[9]

In addition to these main forces moving America in a liberal direction, there have been several important specific forces aiding the growth of liberalism during the last half-century. Perhaps the most important have been the New Deal realignment of the Great Depression (Hamby, 1985; Leuchtenburg, 1983) and the strong leadership role played by certain institutions of government, especially the Supreme Court. The New Deal realignment not only enacted liberal reforms in the 1930s and 1940s, but stimulated later liberal growth by (1) laying a foundation on which later programs could be built, (2) creating a liberal coalition that was able to dominate politics until the 1970s (and on the congressional level until the present), and (3) nurturing a political generation more liberal than its predecessors. In addition, liberalism was also aided by the strong leadership provided by the Supreme Court and other institutions of government and society. Without the intervention of the Supreme Court in such areas as race relations, the rights of the accused, civil liberties, and voting rights, the growth of liberal attitudes would have been stunted (Marshall, 1987). Similarly, the general willingness of presidents to support the Supreme Court decisions with executive powers reenforced the impact of the Court's decisions.

---

[9] Others have written of a gap between ideals and institutions (Huntington, 1981) and norms and experience (McLoughlin, 1978).

Counteracting liberal advances were a series of less general, but still powerful historical forces. The surge in the crime rate starting in the 1960s and of drug-related crime later on triggered public support for social control and punitiveness. Similarly, the stagflation of the 1970s directly counteracted some of the liberalizing tendencies of modernization and reduced public faith in the government's ability to successfully regulate the economy. The governmental failures at economic management and rising tax burdens coupled with its limited success in dealing with some social problems, such as welfare and urban renewal, to lower both confidence in the government and faith in government as a general solution to problems. The disillusionment with government as the solver is illustrated by the mocking shibboleth often cited by Reagan, "I'm from the government and I'm here to help." These problems were instrumental in triggering such liberal reactions as the "mugged by reality" Neoconservatives, the "robbed by the government" tax revolters (Hansen, 1980 and 1983; Sears and Citrin, 1982), and the New Fiscal Populists (Clark and Ferguson, 1983).

Perhaps it was not only liberal failures, but also the unintended consequences of liberal successes, that helped to undermine liberalism. In the area of crime, the extension of the rights of the accused may well have hindered the police in dealing with the rise in crime from the early 1960s to the early 1980s. Similarly, in the area of the economy the protection of employees, consumers, and the environment may have hindered the wealth-producing capacity of the free market. But even more broadly, the great gain in public support for individual choice in areas ranging from abortion, sex roles, and sexual morality to civil liberties may have undermined broad-based, principled support for collective solutions imposed by the government in general.

Cycles in American History

While specific sources of conservative reaction such as crime and stagflation (as well as such events as Vietnam, Watergate, etc.) may offer a sufficient explanation for the stalling of the general liberal movement in the mid-1970s, it is possible that cyclical forces were also at work. This possibility is explored in greater depth elsewhere (Smith, 1984 and 1989) and we will summarize it here.

Various investigators have spotted many different cycles running through American history: business cycles, electoral cycles, religious awakenings, periods of student activism, and so forth. The proposed cycle most relevant to our discussion of liberal-conservative trends is the cycle of reform advanced most prominently by historians Arthur M. Schlesinger, Sr. (1939, 1949, and 1950) and Arthur M. Schlesinger, Jr. (1980, 1984, and 1986). They argue that American history is characterized and shaped by the periodic alternation of periods of liberal reform and conservative retrenchment. During the liberal part of the cycle, the people are activists and back various innovations. Then their reformist zeal weakens and a period of quietude and stand patism prevails. While neither the Schlesingers nor any other proponents of the cycle of reform model have been able to convincingly establish their view of American historical development, they have described a pattern that both matches the shape of the liberal-conservative trend described above (a liberal rise followed by a plateau) and which roughly predicted the timing of both the New Frontier/Great Society and the "Reagan Revolution." Using the observed length of reform cycles through 1931, Schlesinger, Sr. in 1939 predicted a conservative phase starting in 1947 or 1948 and then in 1949 forecasted:

> We may expect the recession from liberalism which began in 1947 to last till 1962, with a possible margin of a year or two in one direction or another. The next conservative epoch will then be due around 1978.

Conclusion

Overall, the post–World War II period has been a time of liberal advance. Liberal trends outnumbered conservative trends by over two-to-one (Duncan, Schuman, and Duncan, 1973; Hamby, 1985; Hoge, 1974; Hoge, Luna, and Miller, 1984; Willits, Bealer, and Crider, 1977). Liberal gains were strongest on such topics as race relations and women's rights that concerned equal rights for all (Gusfeld, 1981; Rokeach and Ball-Rokeach, 1988; Smith and Sheatsley, 1984) and on abortion, civil liberties, and sexual morality that dealt with individual choice (Caplow et al., 1983; Hoge, Luna, and Miller, 1984; McClosky and Brill, 1983; Mueller, 1988). Topics dealing with material concerns and government regulation were mixed in their trends. Responses to calls for more government action were also quite mixed, with the number of trends in opposition to more government edging out trends in favor of more government. In addition, this role of government dimension had little relationship to liberalism/conservatism. Finally, crime was the one topic that consistently showed little or no liberal growth.

Liberal movement slowed appreciably in the mid-1970s and a number of trends, especially in the areas of abortion, civil liberties, crime, and spending and taxes, slowed, stalled, or even, in a few cases, reversed. But the hosannas from the right and wailing from the left over a conservative tide and the Reagan Revolution (Smith, 1982 and 1985a) are both overreactions. On average, liberal momentum and advance ended on the liberal plateau of the mid-1970s, but no general conservative advance occurred.[10]

While the antiliberal problems of crime and economic stagnation offer important particular explanations for many of the changes, a more general cyclical force may also have contributed to the dissipation of reformist momentum. The survey data compiled in this paper are consistent with the cycle of reform hypothesis, but as yet this idea is only suggestive.

References

Abbott, Philip (1981)
"The character of recent American political thought." Canadian Journal of Political Science 14:377–96.
Anagnoson, J. Theodore (1972)
"Ideology in the American electorate, 1956–68." Unpublished paper.
Appelbaum, Richard (1970)
Theories of Social Change. Chicago: Markham Publishing.
Astin, Alexander W., Kenneth C. Green, and William S. Korn (1987)
The American Freshman: Twenty Year Trends. Los Angeles: Higher Education Research Institute.
Bahr, Howard M., Theodore Caplow, and Bruce A. Chadwick (1983)
"Middletown III: Problems of replication, longitudinal measurement, and triangulation." Annual Review of Sociology 9:243–64.
Bell, Roger (1984)
Last among Equals: Hawaiian Statehood and American Politics. Honolulu: University of Hawaii Press.
Boudon, Raymond (1983)
"Why theories of social change fail: Some methodological thoughts." Public Opinion Quarterly 47:143–60.
Brock, Clifton (1962)
Americans For Democratic Action: Its Role In National Politics. Washington, DC: Public Affairs Press.
Burnham, Walter Dean (1980)
"American politics in the 1980s." Dissent 27:149–60.
Cairns, Grace E. (1962)
Philosophies of History: Meeting of East and West in Cycle-Pattern Theories of History. Westport, CT: Greenwood Press.
Caplow, Theodore, Howard M. Bahr, Bruce A. Chadwick, Dwight W. Hoover, Laurence A. Martin, Joseph B. Tamney, and Margaret Holmes Williamson (1983)
All Faithful People: Change and Continuity in Middletown's Religion. Minneapolis: University of Minnesota Press.
Chafetz, Janet Saltzman, and Helen Rose Fuchs Ebaugh (1983)
"Growing conservatism in the United States? An examination of trends in political opinion between 1972 and 1980." Sociological Perspectives 26:275–98.

[10] Most scholarly and semischolarly interpretations have tended to reject the more popularly based "conservative tide" hypothesis. Among its supports are Eismeier, 1982; and Tiryakiian, 1981; giving mixed support is Exter and Barber, 1986; and opponents include Ferguson and Rodger, 1986; Goodman, 1983; Hibbs, 1982 and 1987; Smith and Spinard, 1981; Smith 1982, 1985a, 1985b, and 1985c.

Chong, Dennis, Herbert McClosky, and John Zaller (1983)
"Patterns of support for democratic and capitalist values in the United States." British Journal of Political Science 13:401–40.

Clark, Terry Nichols, and Lorna Crowley Ferguson (1983)
City Money: Political Processes, Fiscal Strain, and Retrenchment. New York: Columbia University Press.

Cochran, Thomas C. (1972)
Social Change in America: The Twentieth Century. New York: Harper and Row.

Conover, Pamela Johnston, and Stanley Feldman (1981)
"The origins and meanings of liberal/conservative self-identification." American Journal of Political Science 25:617–45.

Davis, James A. (1980)
"Conservative weather in a liberalizing climate: Change in selected NORC General Social Survey items, 1972–1978." Social Forces 58:1129–56.

Davis, James A., and Tom W. Smith (1989)
General Social Surveys, 1972–1989: Cumulative Codebook. Chicago: NORC.

Durant, Will, and Ariel Durant (1968)
The Lessons of History. New York: Simon and Schuster.

Eismeier, Theodore J. (1982)
"Drifting right: Changes in public opinion." Journal of Contemporary Studies 5:77–82.

Exter, Thomas, and Frederick Barber (1986)
"The age of conservativism." American Demographics 8 (November): 30–37.

Ferguson, Thomas, and Joel Rogers (1986)
"The myth of America's turn to the right." Atlantic Monthly, May, 43–53.

Duncan, Otis Dudley, Howard Schuman, and Beverly Duncan (1973)
Social Change in a Metropolitan Community. New York: Russell Sage.

Gerber, William (1975)
American Liberalism: Laudable End, Controversial Means. Boston: Twayne.

Gillon, Steven M. (1987)
Politics and Vision: The ADA and American Liberalism, 1947–1985. New York: Oxford University Press.

Glenn, Norval D. (1975)
"Trend studies with available survey data: Opportunities and pitfalls." In Jessie C. Southwick (ed.), Survey Data Trend Analysis: An Index to Repeated Questions in U.S. National Surveys by the Roper Public Opinion Research Center. Williamstown, CT: Roper.

Goodman, John L., Jr. (1983)
Public Opinion During the Reagan Administration. Washington, DC: Urban Institute Press.

Gusfield, Joseph R. (1981)
"Social movements and social change: Perspectives of linearity and fluidity." Research in Social Movements, Conflicts, and Changes 4:317–39.

Hamby, Alonzo L. (1985)
Liberalism and Its Challengers: FDR to Reagan. New York: Oxford University Press.

Hansen, Susan B. (1980)
"Taxes, benefits and public opinion." In Barry S. Rundquist (ed.), Political Benefits. Lexington, MA: D. C. Heath.

——— (1983)
The Politics of Taxation: Revenue Without Representation. New York: Praeger.

Hartz, Louis (1962)
The Liberal Tradition in America: An Interpretation of American Political Thought Since the Revolution. New York: Harcourt Brace.

Hastings, Philip K., and Dean R. Hoge (1981)
"Religious trends among college students, 1948–1979." Social Forces 60:517–31.

Hero, Alfred O., Jr. (1969)
"Liberalism-conservatism revisited: Foreign vs. domestic federal policies, 1937–1967." Public Opinion Quarterly 33:399–408.

Hibbs, Douglas A., Jr. (1982)
"President Reagan's mandate from the 1980 elections: A shift to the right?" American Politics Quarterly 10:387–420.
——— (1987)
The American Political Economy: Macroeconomics and Electoral Politics. Cambridge: Harvard University Press.
Hoge, Dean R. (1974)
Commitment on Campus: Changes in Religion and Values over Five Decades. Philadelphia: The Westminster Press.
Hoge, Dean R., Cynthia L. Luna, and David K. Miller (1981)
"Trends in college students' values between 1952 and 1979: A return of the fifties?" Sociology of Education 54:263–74.
Holton, Robert J. (1981)
"Marxist theories of social change and transition from feudalism to capitalism." Theory and Society 10:833–67.
Huntington, Samuel P. (1981)
American Politics: The Promise of Disharmony. Cambridge: Harvard University Press.
Hyman, Herbert H. (1972)
Secondary Analysis of Sample Surveys: Principles, Procedures, and Potentialities. New York: Wiley.
Kelley, Robert (1977)
"Ideology and political culture from Jefferson to Nixon." American Historical Review 82:531–62.
Kulka, Richard A. (1982)
"Monitoring social change via survey replication: Prospects and pitfalls from a replication of social roles and mental health." Journal of Social Issues 38:17–38.
Ladd, Everett Carll (1983)
"The brittle mandate extended: An in-depth examination of the voters' 1982 verdict." Public Affairs Review 4:23–37.
Leuchtenberg, William E. (1983)
In the Shadow of FDR: From Harry Truman to Ronald Reagan. Ithaca: Cornell University Press.
Libros, Hal (1975)
Hard Core Liberals: A Sociological Analysis of the Philadelphia ADA. Cambridge, MA: Schenkman.
Lowi, Theodore J. (1969)
The End of Liberalism: Ideology, Policy, and the Crisis of Public Authority. New York: W. W. Norton.
Luttbeg, Normal R., and Michael M. Gant (1985)
"Liberal/conservative ideology as a cognitive structure." Public Opinion Quarterly 49:80–93.
McClosky, Herbert, and Alida Brill (1983)
Dimensions of Tolerance: What Americans Believe about Civil Liberties. New York: Russell Sage.
McClosky, Herbert, and John Zaller (1984)
The American Ethos: Public Attitudes toward Capitalism and Democracy. Cambridge, MA: Harvard University Press.
McElvaine, Robert S. (1987)
The End of the Conservative Era: Liberalism After Reagan. New York: Arbor House.
McLoughlin, William G. (1978)
Revivals, Awakenings, and Reform: An Essay on Religious and Social Change in America, 1607–1977. Chicago: University of Chicago Press.
Maddox, William S., and Stuart A. Lilie (1984)
Beyond Liberal and Conservative: Reassessing the Political Spectrum. Washington, DC: Cato Institute.
Mansfield, Harvey C., Jr. (1978)
The Spirit of Liberalism. Cambridge: Harvard University Press.

Marshall, Thomas R. (1987)
"The Supreme Court as an opinion leader: Court decisions and the mass public." American Politics Quarterly 15:147–68.
Martin, Elizabeth (1983)
"Surveys as social indicators: Problems in monitoring trends." In Peter H. Rossi, James D. Wright, and Andy B. Anderson (eds.), Handbook of Survey Research. New York: Academic Press.
Moberg, David O., and Dean R. Hoge (1986)
"Catholic college students' religious and moral attitudes, 1961 to 1982: Effects of the sixties and the seventies." Review of Religious Research 28:104–17.
Mueller, John (1988)
"Trends in political tolerance." Public Opinion Quarterly 52:1–25.
Nisbet, Robert A. (1969)
Social Change and History: Aspects of the Western Theory of Development. London: Oxford University Press.
Ogburn, William Fielding (1966 [1922])
Social Change with Respect to Cultural and Original Nature. New York: Dell.
Pollack, Philip H., III (1983)
"Are there two liberalisms? The partisan and demographic influences on two dimensions of political belief." Sociological Focus 16:227–37.
Robinson, John P., and John Fleishman (1984)
"Ideological trends in American public opinion." Annals 472:50–60.
Rokeach, Milton, and Sandra J. Ball-Rokeach (1988)
"Stability and change in American value priorities, 1968–1981." Paper presented to the American Association for Public Opinion Research, Toronto.
Rotunda, Ronald D. (1986)
The Politics of Language: Liberalism as Word and Sword. Iowa City: University of Iowa Press.
Ryan, Bryce F. (1969)
Social and Cultural Change. New York: Ronald Press.
Schlesinger, Arthur M., Sr. (1939)
"Tides in American politics."Yale Review 29:217–30.
——— (1949)
Paths to the Present. New York: Macmillan.
——— (1950)
The American as Reformer. Cambridge: Harvard University Press.
Schlesinger, Arthur M., Jr. (1980)
"Is liberalism dead?" New York Times Sunday Magazine. 30 March.
——— (1984)
"The election and after." New York Review of Books. 16 August.
——— (1986)
The Cycles of American History. Boston: Houghton Mifflin.
Sears, David O., and Jack Citrin (1982)
Tax Revolt: Something for Nothing in California. Cambridge: Harvard University Press.
Smith, Kathleen Mauer, and William Spinard (1981)
"The popular political mood." Social Policy 11:37–45.
Smith, Tom W. (1976)
"A trend analysis of attitudes towards capital punishment." In James A. Davis (ed.), Studies of Social Change Since 1948. Chicago: NORC.
——— (1980)
"The 75% solution: An analysis of the structure of attitudes on gun control, 1959–1977." Journal of Criminal Law and Criminology 71:300–316.
——— (1982)
"General liberalism and social change in post–World War II America: A summary of trends." Social Indicators Research 10:1–28.
——— (1984)
"Cycles of reform? A summary of trends since World War II." Paper presented to the American Sociological Association.

——— (1985a)
"Atop a liberal plateau?" Research in Urban Policy 1:245–57.
——— (1985b)
"The conservative tide vs. the liberal plateau: Changes in attitudes since World War II." Paper presented to the American Association for Public Opinion Research.
——— (1985c)
"Social change and public opinion since W.W. II." Paper presented to the American Sociological Association.
——— (1988)
"Liberal and conservative trends in the United States since World War II." Paper presented to the American Association for Public Opinion Research.
——— (1989)
"Liberal and conservative trends in the United States since World War II." GSS Social Change Report no. 29. Chicago: NORC.
Smith, Tom W., and Sara P. Crovitz (1988)
Annotated Bibliography of Papers Using the General Social Surveys. 7th ed. Chicago: NORC.
Smith, Tom W., and Paul B. Sheatsley (1984)
"American attitudes toward race relations." Public Opinion 7:14–15, 50–53.
Sorokin, Pitrim A. (1941)
Social and Cultural Dynamics. Vol. 4. New York: Bedminister Press.
Spengler, Oswald (1926–1928)
The Decline of the West. New York: Alfred A. Knopf.
Stinchcombe, Arthur L. (1978)
Theoretical Methods in Social History. New York: Academic Press.
Taylor, D. Garth (1980)
"Procedures for evaluating trends in public opinion." Public Opinion Quarterly 44:86–100.
Thornton, Arland (1985)
"Changing attitudes towards separation and divorce: Causes and consequences." American Journal of Sociology 90:856–72.
Tiryakian, Edward A. (1981)
"Changing values in the United States: Continuity and rupture." Paper presented to the International American Studies Seminar, Mt. Sorak, Korea.
Tominaga, Ken'ichi (1982)
"Typology in the methodological approach to the study of social change." Paper presented to the International Sociological Association.
Willits, Fern K., Robert C. Bealer, and Donald M. Crider (1977)
"Changes in individual attitudes toward traditional morality: A 24-year follow-up survey." Journal of Gerontology 32:681–88.

THE OPINION-POLICY NEXUS IN GERMANY

JOEL E. BROOKS

Abstract This article analyzes the previously unresearched relationship between mass public opinion and public policy in contemporary West Germany. By studying approximately 150 cases over the last decade, the nature of German democracy is revealed in relation to the overall consistency between majority preferences and government action. The opinion-policy nexus is explored in regard to the impact of issue saliency, landslide majorities, different categories of issues (e.g., redistributive, foreign policy), and the partisan composition of the government (i.e., Social Democratic vs. Christian Democratic). In addition, there is a cross-national comparison of results for West Germany with the author's previous research on opinion and policy in the United States, Britain, and France. The findings indicate that (like other nations studied) public opinion and public policy in Germany are *inconsistent* in a majority of instances and that (unlike Britain or France) the partisan composition of the government does not matter vis-à-vis the degree of policy-opinion congruence.

In the period following the establishment of the Federal Republic of Germany, the concern of many analysts revolved around the viability and fragile nature of German liberal democracy. For example, Almond and Verba (1965) implied that West Germany lacked the attributes of an appropriate "civic culture" presumed necessary for stable democracy. Given the legacy of the Nazi era, Ralf Dahrendorf (1966) also raised the issue of German commitment to liberal democratic principles. Recently, however, David Conradt et al. (1983:176) has observed that "the question is no longer whether West Germany will remain a liberal democracy, but rather, what kind of and how much democracy Germany will have" (cf. Smith, 1982; Katzenstein, 1987).

The purpose of this article is to address Conradt's question by applying an underutilized yet potentially enlightening perspective: the opinion-policy nexus. This approach entails empirically testing the

JOEL E. BROOKS is Assistant Professor of Political Science at Illinois State University.

Public Opinion Quarterly 54:508–529 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-04/$2.50

relationship between mass public opinion and governmental policy —establishing a quantifiable and cross-nationally comparable measure of the degree and kind of democracy. Such an approach perceives government responsiveness to mass preferences as a crucial component of a democratic society. As Arend Lijphart explains (1984:1–2), democracy is usually viewed in its representative form—that is, government by the freely elected representatives of the people. However,

> Democracy may be defined not only as government *by* the people but also . . . as government *for* the people—that is, government in accordance with the people's preferences. An ideal democratic government would be one whose actions were always in perfect correspondence with the preferences of all its citizens. Such complete responsiveness in government has never existed and may never be achieved, but it can serve as an ideal to which democratic regimes should aspire. It can also be regarded as the end of a scale on which the degree of democratic responsiveness of different regimes may be measured. (Lijphart, 1984:1–2)

Emphasizing the majoritarian-oriented opinion-policy nexus raises important empirical, analytical, and normative questions. What is the degree of consistency between mass public opinion and public policy in West Germany? How does this consistency rate, compared with other Western democracies? What variables affect this opinion-policy relationship? Is government responsiveness to mass opinion necessarily desirable? The last of these questions (dealing with the normative issue) has been considered in general terms elsewhere (Bachrach, 1967; Pateman, 1970; Butler and Ranney, 1978; Luttbeg, 1981; Sullivan et al., 1981). Within the West German context, public debate has occurred in recent years over the desirability of a number of proposals to increase the influence of average citizens on the political process. These include the direct primary as a device for democratizing the political parties, the introduction of referenda so that issues of significant concern to the public could receive direct popular input, and the extension of citizens' rights to petition and recall elected officials. Such potential (and controversial) reforms are viewed by proponents as mechanisms to increase the impact of mass public opinion on public policy by expanding the participatory and "plebiscitary dimensions" of the democratic constitutional framework in West Germany (Conradt, 1986:242–43; Dalton, 1988:206).

Regardless of whether one accepts or rejects the desirability of a populist congruence between opinion and policy, discerning the degree and type of public opinion–public policy relationship found within different countries can provide a unique insight into the nature of contemporary Western polities. Unfortunately, almost all empirical research on the nexus between opinion and public policy has concentrated solely on the United States (Monroe, 1979; Luttbeg, 1981; Page and

Shapiro, 1983). This article seeks to compensate by providing a comparative cross-national dimension to the opinion-policy nexus. Specifically, this study presents results regarding the relationship between mass opinion and policy for 146 cases spanning the last decade within the Federal Republic of Germany. Potential variations in the opinion-policy nexus are examined in regard to the following variables: (1) extent of majority opinion on an issue (landslide vs. slim majority); (2) issue saliency; (3) types of issues (redistributive vs. distributive/regulatory, foreign/defense policy vs. domestic issues); and (4) partisan composition of the government (SPD/FDP vs. CDU-CSU/FDP). In addition, there is cross-national comparison of results for Germany with previous similar studies, by the author, of opinion and policy in the United States, Britain, and France.

Methodology and Data

The process by which the data were collected is described in detail below. Use was made of published public opinion surveys (Gallup, 1980–81; Noelle-Neumann, 1982; Hastings, 1979–88) done by some of the most reliable institutions in West Germany (Institut für Markt- und Sozialforschung, Institut für Demoskopie Allensbach, Gesellschaft für Marktforschung, EMNID, and the German Gallup affiliate). Each of the surveys met the standards established by Hastings and Hastings for inclusion in their annual *Index to International Public Opinion* (1979:xv). Opinion surveys on national policy-oriented questions from July, 1977, through June, 1987 (the latest ten-year period for which complete data are available), constituted the initial data base. Polls concerning the popularity of current political leaders or parties were automatically eliminated since they did not focus on policy issues, per se. Retrospective opinion questions regarding ex post facto approval or disapproval of previously made governmental decisions were then excluded, since they may not be an accurate reflection of opinion on the issue prior to the announcement of the government's policy. This was done to prevent measuring a phony consensus—that is, where the majority of the public approves the policy only after it is announced.

Surveys on the same specific issue (e.g., support for, or opposition to, a planned natural gas pipeline running from the Soviet Union to West Germany) within the same calendar year were averaged and

---

1. This was done only when the question wording was virtually identical. To do otherwise would allow changes of wording to possibly indicate a false shift of mass opinion. Regarding this and related issues, see the discussion by Yeric and Todd (1989:83–85).

treated as one survey result.[1] This prevented more than one case of opinion-policy consistency (or inconsistency) for any one issue per year. At the same time, issues that appeared over a longer time span (e.g., support for, or opposition to, expansion of the European Community) were considered as separate cases for each year they were polled—reflecting the continued importance of that issue in the eyes of the German polling organizations and the unresolved nature of the issue being considered.[2]

An additional factor is sampling error. Samples of the size used in most of the surveys considered have a 3% margin of error with a confidence level of 95%. As a consequence, all polls indicating a difference between majority and minority opinion falling within this (±3%) range were excluded from the cases analyzed.

As a result of the winnowing process outlined above, 146 issues (cases) remained.[3] For each case, the reported majority opinion on an issue was recorded as either favoring or opposing the policy raised in the survey question. The next step was to characterize government actions in a similar dichotomous fashion on these same specific issues.

In this article, the term *public policy* refers to governmental actions (or inactions) regarding selected issues as indicated by legislative and/or executive decisions. This includes budgetary spending and taxation, passage of domestic legislation, executive orders, patterns of administrative implementation, agreements with foreign governments, and voting records in international or supranational organizations. Excluded from consideration are so-called policy outcomes (e.g., higher unemployment) that may or may not be a direct consequence of a particular governmental action. Also excluded are simple statements of intent not supported by some follow-up governmental action to implement the stated position. The rationale for this exclusion is based on the motto: "Watch what we do, not what we say."

In order to reduce the possibility of error or unintentional bias in determining policy, it was decided to enlist the assistance of two outside specialists in German history and politics. To ensure that their policy judgments were made independent of knowledge regarding how

[2] As Monroe acknowledges (1979:8): "An inherent theoretical problem with this type of research is the question of whether the issues available for study are really the 'important' ones. Unfortunately, there seems to be no purely objective way of answering this. However, we can use the mass public's perception of the salience of various topics." As will be discussed later, of the 146 cases analyzed in this study, only 25 were found to be low in salience (see table 2). For this reason, and because of the high competence of the polling organizations involved, there is good reason to accept these cases as generally representing issues of importance to the mass public.

[3] A complete set of the survey questions utilized is available on request. See the appendix for an alphabetical listing of all 146 cases (including coding decisions).

such evaluations would affect the ultimate findings of this study, the two expert judges had no prior access to the public opinion data. The author's initial characterizations of public policy on the cases under consideration were sent to each of the German specialists. After consultation, a consensus was reached regarding the predominant nature of government action—either generally supportive of, or in opposition to, the policies suggested in the survey questions. At this stage, the author compared the reported majority opinion on each issue (pro or con) with the final (consensus) determination of public policy on that same issue. This yielded the findings in the tables that follow regarding consistency (inconsistency) between mass public opinion and public policy.

An additional element considered in this process was the time-lag factor. How long a period after the expression of public opinion in a survey should one allow before making a determination of policy consistency or incongruence? While there is no obviously right or wrong answer to this question, a 12-month time-lag factor (as adopted by Page and Shapiro, 1983) was utilized.

In order to test for the potential relevance of the type of issue on the opinion-policy nexus in West Germany (discussed in detail in a later section of this article), it was necessary to designate each case according to a dichotomous distinction: "redistributive" versus "nonredistributive." This was based upon a variation of Lowi's typology of public policies (1964 and 1972). For this purpose, all issues involving a potentially progressive (i.e., more equitable) reallocation of economic or political advantage or power in society were categorized as redistributive in nature. Examples would include the following: a proposed increase in West German unemployment compensation programs, extension or strengthening of codetermination provisions, proposals for a flat-rate (vs. earnings related) pension system, an increase in the tax burden on business versus individuals, and measures to integrate and empower migrant workers.

The other distinction utilized regarding type of issue was the domestic–foreign policy dichotomy. In this study, defense issues (e.g., the level of West German military expenditures) were coupled with foreign policy cases (e.g., sanctions against South Africa) when assessing government responsiveness to majority sentiment.

Degree of Consistency

Table 1 presents, inter alia, the overall degree of consistency (inconsistency) between mass public opinion and public policy in West Germany for the ten-year period under consideration. Notice that in al-

Table 1. Overall Opinion-Policy Relationship and Impact of Landslide Majorities

	Landslide		Slim Majority		Overall Relationship	
	%	*N*	%	*N*	%	*N*
Consistency	46	29	40	33	42	62
Inconsistency	54	34	60	50	58	84
All cases	100	63	100	83	100	146

NOTE: Differences between landslide and slim majority cases are not statistically significant at the .05 level using chi-square test.

most six out of ten cases there was *incongruence* between government action and mass preferences. For the minority of instances (42%) manifesting congruence between opinion and policy, the following caveat is essential. A finding of consistency in these cases does *not* necessarily establish a causal relationship (i.e., a presumed "linkage") between mass opinion and resulting governmental policy (Key, 1961; Sullivan, 1974; Weissberg, 1976; Luttbeg, 1981).

An alternative explanation, referred to as the "counterfeit consensus" school of thought, views mass public opinion as nonautonomous in nature (Brooks, 1985:250–51). From this perspective, opinion is depicted as being largely a creation of the ruling elites. First, it is contended that governmental elites can mold mass attitudes through the "phony consensus" device of rallying patriotic fervor in times of perceived crisis to create artificially high support for government policy (i.e., the "rally-round-the-flag" syndrome). Note, however, that an effort is made to exclude most cases of this type (i.e., those involving ex post facto approval of policy) from consideration in this study. Second, it is argued that mass opinion tends to reflect the dominant values of the ruling business elite because of its influence over the modes of opinion formation and the socialization (legitimation) process (Bottomore, 1964; Edelman, 1964; Miliband, 1969; Marsh, 1983). As a consequence, there would be a preordained consistency between opinion and policy. Such a consensus would be "counterfeit" in that it would not derive from genuinely independent mass opinion but would be the result of a manipulative process (Hewitt, 1974). Therefore, if this argument is even partially correct, the percentage of cases suggesting congruence between opinion and government action in West Germany may actually exaggerate the real degree of majoritarian democracy that exists. In essence, the true consistency rate may be even lower than the 42% indicated in table 1. On the other hand, the ability of the ruling elites to manipulate mass opinion (creating an artificial consensus) most of the time is open to doubt given the lack of high correspondence between mass preferences and policy in a majority of the cases analyzed.

Comparative Assessment

IMPACT OF INSTITUTIONAL STRUCTURES

Studies of the policy performance of Western democracies in general (as well as West Germany in particular) often assert the importance of major institutional arrangements regarding the distribution and exercise of governmental power—for example, federal versus unitary, parliamentary as opposed to presidential structures (Katzenstein, 1987; Conradt, 1986; Lijphart, 1984; Heidenheimer, Heclo, and Adams, 1983; Powell, 1982). Do such variations in institutional structures matter vis-à-vis the opinion-policy nexus? A tentative answer can be provided by comparing the results for West Germany (a parliamentary, federal state) with those obtained by the author in previous studies of Britain (a parliamentary, unitary state), the United States (a presidential, separation-of-powers, federal system), and France (a hybrid presidential, unitary nation). For none of the four nations does opinion-policy consistency even reach the 50% level. In fact, the congruence rates range from 44% for Britain to 40% for France (Brooks, 1985 and 1987)—extremely similar to the findings for West Germany.[4] Thus, obvious institutional differences between Britain, France, Germany, and the United States are apparently less important than the common factors among these nations affecting the opinion-policy nexus. This raises profound questions about the "democratic nature" (as previously defined by Lijphart) of developed capitalist societies in which majority preferences are thwarted on a majority of policy issues. The emphasis of the remainder of this study is to focus specifically on potential variations in the opinion-policy nexus based on empirical analysis of the variables discussed earlier.

---

[4] My previous research has obtained different results for the overall American opinion-policy nexus than found by Monroe (1979) or Page and Shapiro (1983). This is probably due to differences in methodology and interpretation of findings. For example, Page and Shapiro used what they term a "change/covariation design" (1983:176–77) that establishes a less stringent standard for consistency between opinion and policy. They identified instances in which there was a "significant change" (i.e., 6 percentage points or more) in opinion from one survey to another regarding the same issue. They then measured policy outputs to see whether policy moved in the same direction as opinion. Page and Shapiro (1983:177) found congruent changes in policy 43% of the time (vs. a finding of 41% congruence in my research). Thus, there was an initial similarity of results. However, Page and Shapiro decided to reinterpret cases where opinion changed but policy did not (1983:178). As a consequence, they excluded most such cases from their analysis—artificially inflating the apparent degree of opinion-policy consistency.

IMPACT OF LANDSLIDE MAJORITIES

Despite the results reported above, it is conceivable that an overwhelming degree of majority opinion (i.e., a landslide majority) regarding an issue might carry greater weight than a slim majority or plurality. Thus, one can hypothesize: consistency between mass public opinion and public policy is higher for landslide majority than for slim majority (or plurality) cases. For the purposes of this article, all cases involving at least 60% majority opinion on an issue are designated as "landslide majority."

As table 1 indicates, there is only a minor and statistically insignificant difference between the two categories of cases. Contrary to what was hypothesized, it appears that the extent of majority opinion expressed in opposition to, or in support of, a particular issue has not been a decisive factor in the decision-making process in West Germany.

On a comparative basis, similar findings exist. The consistency rate for landslide cases ranges from 47% for Britain to 42% for the United States (Brooks, 1985:256). This is not significantly different from the rate for slim majority cases in these other nations. Thus, a landslide majority fails to achieve a greater degree of success in affecting policy than smaller majorities.

ISSUE SALIENCE

Clearly, not all issues are of equal importance to the public. On issues that the public apparently cares a great deal about, there might be greater pressure to respond on the part of the government. Therefore, one could plausibly hypothesize that policy-opinion congruence would be higher for salient versus nonsalient issues. If this were the case, such a factor would affect the interpretation of results provided so far.

Since most public opinion surveys do not simultaneously measure salience as well as direction of opinion on specific questions, a surrogate for a direct indicator of salience is necessary. As Page and Shapiro (1983:181) suggest:

> Salience can be measured more directly, although still not perfectly, by the proportion of respondents answering "don't know" or "no opinion" to survey questions. When the proportion of don't knows is relatively low—that is, when more people are willing to offer a preference—it is a

sign of more public interest and attention and perhaps also stronger, more intensely held opinions.

Following the practice estabiished by Page and Shapiro, issues for which the percentage of "don't knows" or "no opinions" equaled 5% or less were categorized as high salience, 6–14% as moderate salience, and 15% or more as low salience. The results are presented in table 2. Slightly over one-quarter of all issues fell within the high salience category. However, the degree of congruence between mass public opinion and public policy was no greater for these cases than for the low salience issues. Contrary to conventional wisdom, but consistent with the evidence of substantial lack of democratic responsiveness in West Germany provided in this article, high salience does not lead to a greater degree of correspondence between opinion and policy.

REDISTRIBUTIVE ISSUES

Another way to approach the opinion-policy nexus is to differentiate between types of issues. As noted earlier, one can adapt Lowi's typology of policies—distinguishing between redistributive and nonredistributive cases (the latter category combining what Lowi calls "regulatory" and "distributive"). To reiterate, redistributive issues are defined in this study as proposals to reallocate income, property, political rights, or some other related value among broad groups or classes. This necessarily means that there will be winners and losers. This process often will be marked by class conflict and will be cast in ideological terms. It can be argued that on redistributive types of issues (i.e., those that potentially threaten the economic elites' continued predominance), the degree of consistency between opinion and policy ought to be even lower than that which generally occurs. Since these issues challenge those forces that benefit from existing inequalities in economic and political power, it can be expected that in these cases, more than others, attempts will be made to frustrate the popular will. This can be anticipated for developed capitalist democracies in general (Marsh, 1983)—including West Germany (Webber, 1983).

As a consequence of the argument above, one can hypothesize: consistency between mass public opinion and public policy is lower for redistributive than for nonredistributive issues. The results in table 3 confirm the hypothesis. For the redistributive cases considered, consistency scores are substantially lower (19%) than for nonredistributive cases (52%). This finding is compatible with the views expressed by Lindblom (1977) and Marsh (1983:4) that the economic elite in Western societies primarily "exercises veto power on those issues which most crucially affect it. . . . In contrast, on secondary [i.e., nonredistribu-

Table 2. Issue Salience

	Low		Moderate		High		Total	
	%	*N*	%	*N*	%	*N*	%	*N*
Consistency	40	10	44	36	41	16	42	62
Inconsistency	60	15	56	46	59	23	58	84
All cases	17	25	56	82	27	39	100	146

NOTE: Differences are not statistically significant.

Table 3. Redistributive Issues

	Redistributive		Nonredistributive	
	%	*N*	%	*N*
Consistency	19	8	52	54
Inconsistency	81	34	48	50
All cases	100	42	100	104

NOTE: Differences are statistically significant at the .05 level using the chi-square test.

tive] issues, other groups may occupy a privileged position and . . . business itself will disagree amongst itself. . . . So while business enjoys a veto on the key issues there is pluralism on secondary issues."

For Britain, the United States, and France, the degree of consistency between opinion and policy for redistributive issues (while lower than that for nonredistributive cases) averaged twice the rate for West Germany (Brooks, 1985 and 1987). From this evidence, one might argue that German elites appear to be even more powerful in their ability to thwart redistributive policies incompatible with their interests than are elites in the other Western nations under consideration. The extent to which these results are a function of the particular development and nature of the West German political party system is examined in detail in the last section of this article.

FOREIGN POLICY ISSUES

Another distinction between different categories of issues involves foreign versus domestic cases. There is a frequently espoused, though largely unproven, assumption in foreign policy and public opinion literature that foreign policy is less susceptible to pressures of mass opinion (Rosenau, 1961; Macridis, 1972). This is supposedly true due to the presumed lower salience of many foreign policy issues for the masses and greater interest group pressure vis-à-vis domestic concerns. While such ideas have been presented mainly in regard to the United States, they have been suggested in relation to West Germany as well (Deutsch and Edinger, 1973; Merritt and Merritt, 1980). As a consequence, one can hypothesize: consistency between mass public opinion and public policy is lower for foreign policy than for domestic issues.

The results in table 4 show the degree of congruence is actually higher for foreign policy than domestic issues. This pattern is found for France, Britain, and the United States as well (Brooks, 1985 and 1987). One possible explanation is that the higher consistency rate for foreign policy cases is a reflection of the "counterfeit consensus" process described earlier. It can be argued that an artificial consensus is more easily obtained in foreign policy cases given the "rally-round-the-flag" syndrome, the ability of governmental elites to control information more easily regarding such issues, and the relatively lower level of public knowledge usually associated with foreign affairs. Alternatively, the findings can be explained in light of the discussion above regarding lower congruence between opinion and policy for redistributive cases. Since virtually all such issues are also domestic in nature, lower consistency scores result for domestic as opposed to foreign policy/defense cases (given the large number of redistributive issues in the former category).

PARTISAN COMPOSITION OF THE GOVERNMENT

German political scientist Kurt Sontheimer has referred to the Bonn Republic as a "party state." In his words (1973:95): ". . . all political decisions in the Federal Republic are made by the parties and their representatives. There are no political decisions of importance in the German democracy which have not been brought to the parties, prepared by them and finally taken by them." In fact, the West German Basic Law (unlike most other democratic constitutions) specifically mentions political parties in detail and assigns them prominent functions (in article 21). If the role of political parties in the policy process is as great as suggested, then the following question is pertinent: Do changes in the partisan composition of the West German government make a difference vis-à-vis the opinion-policy nexus?

Two rival hypotheses will be presented. On the one hand, it can be hypothesized that the ideology of the German Social Democrats (SPD)

Table 4. Domestic Versus Foreign Policy Issues

	Foreign Policy		Domestic Cases	
	%	*N*	%	*N*
Consistency	56	28	35	34
Inconsistency	44	22	65	62
All cases	100	50	100	96

Note: Differences are statistically significant at the .05 level using the chi-square test.

could lead to greater consistency between mass public opinion and policy than under Christian Democratic (CDU) dominated governments. This would be due primarily to differences resulting from decisions on *redistributive* issues. It can be argued that under conditions of Social Democratic rule, the supposed representatives of the working class (and the alleged opponents of the economic elite) have attained control of the mechanisms of government. As a consequence, action regarding redistributive cases would be more responsive to mass interests and preferences. As Russell Dalton (1988:134–35) argues:

> The SPD emerged from the socialist working-class movement and still consistently represents working-class interests. Although West German labor unions no longer have institutional ties to the SPD, the relationship nevertheless remains close. For instance, over half of the SPD Bundestag deputies belong to a labor union; labor union members constitute the SPD's core voters; and the campaign support of labor unions is crucial to the SPD's electoral chances. Thus, the SPD's control of the national government from 1969 until 1982 resulted in a series of policies designed to improve the situation of the average worker: expanded social programs, retirement benefits, health-care insurance, and worker self-management (*Mitbestimmung*).

Alternatively, it can be hypothesized that there is no significant difference between SPD versus CDU dominated administrations in relation to the opinion-policy nexus. For example, as Robert Livingston has observed (1987:iv):

> When Helmut Kohl's Christian Democrats came to power in 1982, they at first gave the impression that they might follow free market policies like those of Kohl's fellow conservatives, Ronald Reagan and Margaret Thatcher. A "U-turn" (*Wende*) was promised. . . . [However] the Christian Democrats have hardly tried to move against the West German welfare state, other than imposing some fiscal discipline. They have not reduced its subsidies, its regulations, or its social spending.

In regard to the Social Democrats, there are at least three major reasons for expecting little distinction from CDU policy. First, since the Bad Godesberg program of 1959, it can be argued that the SPD has been transformed from a traditional "workers' party" to a more moderate "people's party" (*Volkspartei*). As Conradt (cf. Braunthal, 1983; Pappi, 1984; Livingston, 1986; Paterson, 1986) explains, the SPD has been

> . . . concerned above all with embracing the middle, that is, appealing for middle-class support by stressing its allegiance to the free market economy and the Western alliance. . . . Government power had come largely through the party's conscious move into the center of the political spectrum. . . . The left argued that the party . . . had taken the easy route to political power by being content simply to represent diverse social groupings and classes without changing the power relationships between them. (1986:95–96)

Second, the Free Democratic party has held the balance of power in the Bundestag continuously since 1969. This has been as a junior partner in coalition with the Social Democratic Party until switching to an alliance with the Christian Democrats in the fall of 1982. The further moderating influence of the FDP on the SPD has been suggested by many analysts—especially in regard to economic issues (George, 1985:181; Pappi, 1984:23; Webber, 1983:75). Third, some SPD proposals have been altered by the Christian Democrats through their position in the Bundesrat (Conradt, 1986:241). As a consequence, potential redistributive actions by the SPD could have been blocked or modified.

Table 5 presents the results regarding the impact of the partisan composition of the government on the opinion-policy nexus. On an overall basis, the findings show no statistically significant distinction exists for SPD/FDP versus CDU-CSU/FDP governments. Regardless of the party in power, majority sentiment is thwarted in a majority of the cases. What about landslide majorities? Table 5 indicates only minor differences in responsiveness between the parties. Such differences are not statistically significant. In relation to types of issues, the table shows no discernible difference between political parties regarding the degree of consistency for foreign policy cases.

The pattern continues in regard to the findings for redistributive issues. Note that table 5 indicates there is a virtually identical high rate of inconsistency between opinion and policy regardless of the partisan composition of the government. Attainment of political power by the West German Social Democratic Party does *not* appear to have an impact upon the very issues affecting its supposed ideological raison d'être. As Katzenstein has argued (1987:4): "West Germany's SPD had in fact been quite conservative in the 1970s, eschewing both French-style socialism and Swedish-style social democracy."

Table 5. Partisan Composition of the Government

	CDU-CSU/ FDP		SPD/FDP	
	%	*N*	%	*N*
Overall pattern				
Consistency	46	32	39	30
Inconsistency	54	38	61	46
All cases	100	70	100	76
Landslide cases				
Consistency	48	15	44	14
Inconsistency	52	16	56	18
Total	100	31	100	32
Foreign policy cases				
Consistency	57	16	55	12
Inconsistency	43	12	45	10
Total	100	28	100	22
Redistributive cases				
Consistency	18	4	20	4
Inconsistency	82	18	80	16
Total	100	22	100	20

NOTE: None of the differences are statistically significant at the .05 level using the chi-square test.

On a comparative basis, the results in table 5 are *not* consistent with previous research on the opinion-policy nexus. For both Britain and France, congruence between opinion and policy for redistributive issues was significantly higher under British Labour and French Socialist Party rule than during "bourgeois" governments (Brooks, 1985:257 and 1987:471). Thus, from the standpoint of the opinion-policy relationship, the partisan composition of the government matters (vis-à-vis redistributive issues) in France and Britain but not in West Germany. The results in regard to Germany are compatible with the observation of Douglas Webber (1983:80):

> Where nevertheless "socialist tendencies" threatened to get the upper hand in the SPD (which, especially under Schmidt, was very seldom), these could be neutralised either by the coalition partner, the FDP, or in the Bundesrat by the CDU-CSU majority. The structure of the West German political system militates towards the formation of an effective "all-party Government," which offers organized interests a multiplicity of arenas in which they can bring their influence to bear and which enabled capital and its organisations often to bring theirs to bear even when the SPD was immune to their pressure.

Conclusion

Because the findings discussed in this article are based on a relatively short time-frame (1977–87) encompassing only one chancellor from each of the two major parties, the conclusions regarding the opinion-policy nexus in West Germany should be viewed as tentative. Nevertheless, the results suggest the following generalizations: (1) public opinion and public policy are inconsistent in a majority of cases in Germany—similar to Britain, France, and the United States; (2) landslide majorities in favor of a particular policy are no more likely to be heeded than slim majorities; (3) inconsistency between opinion and policy rises to overwhelming levels (approximately four out of five cases in West Germany) when redistributive issues are concerned; (4) unlike Britain or France, the partisan composition of the German government makes no statistically significant difference in regard to overall consistency rates or in relation to the degree of policy-opinion congruence for redistributive issues; and (5) variations in institutional structures between Germany and the other countries considered in this study do not appear to make a difference at the macrolevel of *overall* congruence between opinion and policy. However, a caveat is in order regarding this last point. The German electoral system (utilizing proportional representation for half the seats in the Bundestag) makes moderating coalitions (with the small FDP) more likely than is the case in Britain or France. In addition, the federal structure in West Germany has often permitted a conservative CDU-CSU majority in the Bundesrat the potential power to modify or even block proposals by a national SPD government. Both of these institutional factors play a role in the lower rate of opinion-policy consistency (reported earlier) for redistributive proposals in West Germany as compared to Britain or France and contribute to the lack of impact of the partisan composition of the government vis-à-vis redistributive issues. Thus, at the microlevel of congruence between opinion and policy for *redistributive* cases, German institutional structures do appear to matter by constraining the more progressive tendencies of the SPD.

Much has been written from Anglo-American, West German, and other perspectives concerning the status of contemporary German democracy. In summarizing the current literature, Jesse (1986:154) contends that "the Federal Republic is a 'normal' democracy just like others." What do the conclusions of this study tell us in this regard? In response to Conradt's question posed at the beginning of this article, what kind of and how much democracy does West Germany have? On one hand, the Federal Republic is "abnormal" (in comparison to Britain and France) in regard to the lack of impact on the opinion-policy relationship resulting from changes in the partisan composition of the government and in relation to its higher degree of inconsistency for redistributive cases. On the other hand, based upon comparative findings cited earlier, West Germany is indeed "normal" in its extent of *overall* inconsistency between public opinion and public policy (occurring in a majority of instances). While the Federal Republic may have evolved into a typical *liberal* democracy (involving free speech, competitive elections for public office, etc.), based upon the yardstick of the opinion-policy nexus applied in this study, West Germany also is "normal" in its failure to become a *majoritarian* democracy.

Appendix

Table A1 provides an alphabetical listing of all 146 cases (including coding decisions) utilized in this study. The procedure for selection of cases is discussed in detail in the Methodology and Data section of this article. The following coding rules (established by the author based upon reasons outlined in the Methodology section) were used by the two expert judges in characterizing public policy: (1) apply a 12-month time-lag factor between the date of the opinion survey and a determination of policy on the issue raised in the survey question; (2) ignore government statements of intended policy, per se, not supported by concrete action to begin implementation; (3) examples of governmental action include legislative and/or executive decisions regarding budgetary spending, taxation, passage of domestic legislation, executive orders, patterns of administrative implementation, written agreements with foreign governments, and voting records in international or supranational organizations; (4) ignore policy "outcomes" (e.g., unemployment or inflation rates) that may or may not be a direct consequence of a particular government decision; and (5) in cases involving *partial* approval or implementation of a proposed policy, make an informed assessment as to whether the government action was closer to being in support of or in opposition to the policy suggested in the survey question. It is precisely because such a determination is not always clear-cut and involves expertise plus judgment that two outside specialists in German modern history and politics were utilized. Note, however, that a consensus ultimately was reached among the expert judges and the author regarding the characterization of all cases.

In table A1 the letter *s* in the appropriate column indicates opinion and/or policy is supportive of the proposal raised in the survey question; the letter *o* refers to opposition toward the proposal. Cases of opinion-policy consistency, therefore, are those with the same letter in both the opinion and policy columns. An *x* in the column labeled "red" refers to a redistributive case (as previously defined). The column labeled "year" refers to when the survey of opinion occurred. The abbreviation "LM" stands for "landslide majority."

Table A1. List of 146 Cases (in alphabetical order)

Issue	Year	Opin.	Policy	LM	Red.
Abortion, liberalization	82	O	O	X	
	83	O	O	X	
	84	O	O		
	85	O	O		
Bundesrat, reduction of power	83	S	O		X
	84	S	O		X
Bureaucracy, decrease salary	80	S	O		
	81	S	O		
Bureaucracy, decrease size	77	S	O	X	
	78	S	O	X	
Conscription of women into noncombat role	83	O	O		
	84	O	O		
Corporations, support for failing firms	81	S	S		X
	82	S	S		X
Criminals, increase punishment of violent offenders	86	S	O	X	
	87	S	O	X	
Death penalty for murder	77	S	O		
	78	S	O		
Defense spending increased above rate of inflation	77	S	O		
	78	S	O		
	80	S	O		
	81	S	O		
	83	S	S		
	84	S	S	X	
Defense spending, reduce and use savings for social programs	86	S	O	X	X
	87	S	O	X	X
Disabled, increase expenditures for	81	S	O		
	82	S	O		
Drugs, crackdown on availability and increase penalties	78	S	O	X	
	79	S	O	X	
East bloc, establish better ties	78	S	S	X	
	79	S	S	X	
Economic aid, increase to third world	79	S	O	X	
	80	S	O	X	
	81	O	O		
	82	O	O		
	83	O	O		
	84	O	O		
Economy, greater role for government management	81	S	O	X	X
	82	S	O	X	X
Energy, establish department of	78	S	S	X	
	79	S	S	X	

Table A1. (*Continued*)

Issue	Year	Opin.	Policy	LM	Red.
Environment, strengthen laws to protect	81	S	S	X	
	82	S	S	X	
	83	S	S	X	
	84	S	S	X	
Equalization payments to needy regions of country	81	S	S	X	X
	82	S	S	X	X
European Community, accelerate unification	77	S	O	X	
	78	S	S	X	
	79	S	S		
	80	S	S		
	86	S	S		
	87	S	S		
European Community, equalization payments to poorer members	79	O	S	X	X
	80	O	S	X	X
	81	O	S		X
	82	O	S		X
	83	O	S		X
	84	O	S		X
	85	O	S		X
	86	O	S		X
European Community, expansion to admit Portugal	82	S	O		
	83	S	O	X	
	84	S	S	X	
	85	S	S	X	
European Community, expansion to admit Spain	77	S	O		
	78	S	O		
	82	S	O		
	85	S	S		
European Parliament, increase power over other EC institutions	83	S	O	X	X
	84	S	O	X	X
Federalism, more decentralization	81	S	O		
	82	S	O		
Foreign workers, citizenship to their children born in West Germany	79	S	O		X
	80	S	O		X
Foreign workers, encourage return to native country	82	S	O		
	83	S	O		
	84	S	S	X	
	85	S	S		
Foreign workers, social benefits fully available	86	S	O		X
	87	S	O		X
Fortieth anniversary of end of World War II, special commemoration	84	O	S		
	85	O	S		

Table A1. (*Continued*)

Issue	Year	Opin.	Policy	LM	Red.
Immunity from prosecution of Bundestag members	77	O	S	X	
	78	O	S	X	
Inequality of income, greater effort to reduce	81	S	O	X	X
	82	S	O	X	X
	83	S	O	X	X
	84	S	O	X	X
NATO, maintain current role	81	S	S		
	82	S	S		
	83	S	S	X	
	84	S	S	X	
Natural gas pipeline with the USSR	82	S	S	X	
	83	S	S	X	
Nazi war criminals still prosecuted	77	S	O		
	79	S	O		
	86	S	O		
	87	S	O		
No-confidence vote as method of government coming to power	83	O	S	X	
	84	O	S	X	
Nuclear energy, continued development	77	O	S		
	78	O	S		
	79	S	S		
	80	S	S		
	81	S	S	X	
	82	S	S	X	
	83	S	S		
	84	S	S		
	85	S	S		
	86	S	S		
Olympic boycott of 1980 Moscow games	79	S	S		
	80	S	S		
Pershing II and cruse missiles deployed	81	O	S		
	82	O	S		
	83	O	S		
	84	O	S		
President, direct election	78	S	O		X
	79	S	O		X
Protectionist trade-policy to aid domestic firms	82	S	O		
	83	S	O		
Public ownership of industries expanded	81	O	O		X
	82	O	O		X
	86	O	O		X
	87	O	O		X

Table A1. (*Continued*)

Issue	Year	Opin.	Policy	LM	Red.
Seat belt, mandatory use	79	S	S		
	80	S	S		
Student aid not to be cut	84	S	O	X	X
	85	S	O	X	X
Taxes on liquor and tobacco increased to support social programs	78	S	O	X	X
	79	S	O	X	X
Television, establishment of additional private stations	83	S	S	X	
	84	S	S	X	
Terrorism, reduction of civil liberties to fight	78	S	O		
	79	S	O		
Terrorism, tougher penalties for acts of	83	S	S	X	
	84	S	S	X	
Unemployment, government enact programs to reduce	78	S	O		X
	79	S	O		X
	83	S	O	X	X
	84	S	O	X	X
War toys, prohibit sale of	78	S	O	X	
	79	S	O	X	
West Berlin, maintain status	78	S	S	X	
	79	S	S	X	

References

Almond, G., and S. Verba (1965)
The Civic Culture: Political Attitudes and Democracy in Five Nations. Boston: Little, Brown.
Bachrach, P. (1967)
The Theory of Democratic Elitism. Boston: Little, Brown.
Bottomore, T. (1964)
Elites and Society. London: Penguin.
Braunthal, G. (1983)
The West German Social Democrats 1969–82: A Profile of a Party in Power. Boulder: Westview Press.
Brooks, J. (1985)
"Democratic frustration in the Anglo-American polities: A quantification of inconsistency between mass public opinion and public policy." Western Political Quarterly 38:250–61.
——— (1987)
"The opinion-policy nexus in France: Do institutions and ideology make a difference?" Journal of Politics 49:465–80.
Butler, D., and A. Ranney (1978)
Referendums: A Comparative Study of Practice and Theory. Washington, DC: American Enterprise Institute.

Conradt, D. (1986)
The German Polity. White Plains, NY: Longman.
Conradt, D., D. Jacobs, B. G. Peters, and W. Safran (1983)
Comparative Politics: United Kingdom, France, Germany, Soviet Union. Chatham, NJ: Chatham House.
Dahrendorf, R. (1966)
Society and Democracy in Germany. Garden City, NY: Doubleday.
Dalton, R. (1988)
Citizen Politics in Western Democracies: Public Opinion and Political Parties in the United States, Great Britain, West Germany, and France. Chatham, NJ: Chatham House.
Deutsch, K., and L. Edinger (1973)
Germany Rejoins the Powers: Mass Opinion, Interest Groups and Elites in Contemporary German Foreign Policy. New York: Octagon Press.
Edelman, M. (1964)
The Symbolic Uses of Politics. Urbana, IL: University of Illinois Press.
Gallup, G. (1980–81)
The International Gallup Polls: Public Opinion in 1978 and 1979. Wilmington, DE: Scholarly Resources.
George, S. (1985)
Politics and Policy in the European Community. New York: Oxford University Press.
Hastings, E., and P. Hastings (1979–88)
Index to International Public Opinion: 1978–87. Westport, CT: Greenwood Press.
Heidenheimer, A., H. Heclo, and C. Adams (1983)
Comparative Public Policy: The Politics of Social Choice in Europe and America. New York: St. Martin's.
Hewitt, C. (1974)
"Policy-making in post-war Britain: A nation-level test of elitist and pluralist hypotheses." British Journal of Political Science 4:187–216.
Jesse, E. (1986)
"The Federal Republic of Germany: From Anglo-American, West German and French perspectives." West European Politics 9:153–58.
Katzenstein, P. (1987)
Policy and Politics in West Germany. Philadelphia: Temple University Press.
Key, V. O. (1961)
Public Opinion and American Democracy. New York: Knopf.
Lijphart, A. (1984)
Democracies: Patterns of Majoritarian and Consensus Government in Twenty-One Countries. New Haven, CT: Yale University Press.
Lindblom, C. (1977)
Politics and Markets: The World's Political-Economic Systems. New York: Basic Books.
Livingston, R. (1986)
West German Political Parties. Baltimore: Johns Hopkins University Press.
——— (1987)
"Foreword." In N. Walter (ed.), West Germany's Economy. Baltimore: Johns Hopkins University Press.
Lowi, T. (1964)
"American business, public policy, case studies, and political theory." World Politics 16:677–715.
——— (1972)
"Four systems of policy, politics, and choice." Public Administration Review 32:298–310.
Luttbeg, N. (1981)
Public Opinion and Public Policy. Homewood, IL: Dorsey Press.
Macridis, R. (1972)
Foreign Policy in World Politics. Englewood Cliffs, NJ: Prentice-Hall.

Marsh, D. (1983)
"Interest group activity and structural power: Lindblom's 'Politics and Markets.' " West European Politics 6:3–13.
Merritt, A., and R. Merritt (1980)
Public Opinion in Semisovereign Germany. Urbana, IL: University of Illinois Press.
Miliband, R. (1969)
The State in Capitalist Society. London: Basic Books.
Monroe, A. (1979)
"Consistency between public preferences and national policy decisions." American Politics Quarterly 7:3–19.
Noelle-Neumann, E. (1982)
The Germans: Public Opinion Polls. Westport, CT: Greenwood Press.
Page, B., and R. Shapiro (1983)
"Effects of public opinion on policy." American Political Science Review 77:175–90.
Pappi, F. (1984)
"The West German party system." West European Politics 7:7–26.
Pateman, C. (1970)
Participation and Democratic Theory. London: Cambridge University Press.
Paterson, W. (1986)
"The German Social Democratic Party." In W. Paterson and A. Thomas (eds.), The Future of Social Democracy. New York: Oxford University Press.
Powell, G. B. (1982)
Contemporary Democracies: Participation, Stability, and Violence. Cambridge, MA: Harvard University Press.
Rosenau, J. (1961)
Public Opinion and Foreign Policy. New York: Random House.
Smith, G. (1982)
Democracy in Western Germany. London: Heinemann.
Sontheimer, K. (1973)
The Government and Politics of West Germany. New York: Praeger.
Sullivan, J. (1974)
"Linkage models of the political system." In A. Wilcox (ed.), Public Opinion and Political Attitudes. New York: Wiley.
Sullivan, J., G. Marcus, S. Feldman, and J. Piereson (1981)
"The sources of political tolerance: A multivariate analysis." American Political Science Review 75:92–104.
Webber, D. (1983)
"A relationship of 'critical partnership'? Capital and the social-liberal coalition in West Germany." West European Politics 6:61–86.
Weissberg, R. (1976)
Public Opinion and Popular Government. Englewood Cliffs, NJ: Prentice-Hall
Yeric, J., and J. Todd (1989)
Public Opinion: The Visible Politics. Itasca, IL: F. E. Peacock.

THE BERELSON PARADOX RECONSIDERED

INTENTION-BEHAVIOR CHANGERS IN U.S. AND SWEDISH ELECTION CAMPAIGNS

DONALD GRANBERG and
SÖREN HOLMBERG

Abstract The Berelson paradox is that in democracies, the aggregate requirement for adaptability is provided by the least qualified voters, i.e., volatile voters tend to be less knowledgeable and less involved in politics. Analysis of data from ten U.S. and seven Swedish national election studies pointed to a significant interaction effect. The early findings regarding interest and knowledge were consistently replicated in the U.S., but not in Sweden. In Sweden, intention-behavior changers were not likely to be low in interest or knowledge, and interested nonpartisans were overrepresented among the intention-behavior changers.

A protean democratic system must have some people within it whose behavior is changeable. If everyone consistently voted or did not, without individual variation across time, and if all voters voted on each occasion for the same party, there would be little incentive for a party to adapt its orientation to changing circumstances. The only sources of change would then be attrition through emigration and death or people otherwise losing the right to vote, and extension of the franchise, people coming of voting age, immigration, or otherwise gaining the right to vote. It has been consistently found, however, that such

---

DONALD GRANBERG is currently Waernska Visiting Professor of Political Science at the University of Goteborg, on leave from his position as Professor of Sociology and Research Associate at the Center for Research in Social Behavior at the University of Missouri–Columbia. SÖREN HOLMBERG is Docent in the Department of Political Science at the University of Göteborg and Director of the National Election Studies Project in Sweden. The authors wish to thank Mikael Gilljam, Peter Esaiasson, Iris Alfredsson, Martin Bennulf, Patricia Shanks, Sandy Juergensmeyer, LeeAnn Debo, and Mark Henry for their advice and assistance. The data were obtained through the Inter-University Consortium for Political and Social Research at the University of Michigan and the Svensk Samhällsvetenskaplig Datatjänst at the University of Göteborg. The authors are solely responsible for the analyses and interpretation.

Public Opinion Quarterly 54:530–550 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-04/$2.50

movement in and out of the electorate typically does *not* provide the variation in voting tendencies necessary for a system to be adaptive (Holmberg and Gilljam, 1987; van der Eijk and Niemöller, 1983).

Thus, those continuing in the electorate across time must have some potential for change in order for a democratic system to adapt, although we can not specify an optimal level of volatility within an electorate. In analyzing volatility, Campbell et al. (1960) differentiated between turnout variability (percentage who vote in some but not in every election) and partisan variability (percentage of voters who do not vote for the same party in each election). It is also worthwhile to differentiate across-election volatility from within-campaign volatility (Holmberg, 1981).

At the individual level of analysis, questions relating to voter volatility can be most effectively addressed with panel data. This was done first in Erie County, Ohio, in 1940 and Elmira, New York, in 1948, as reported in the classics, *The People's Choice* (Lazarsfeld, Berelson, and Gaudet, 1944) and *Voting* (Berelson, Lazarsfeld, and McPhee, 1954). It was out of these studies that an intriguing paradox emerged: ". . . the voters least admirable when measured against individual requirements contribute most when measured against the aggregate requirement for flexibility" (1954:316). Stated bluntly, people who were volatile across elections or within an election campaign tended to be disproportionately apathetic and ignorant. What flexibility the system had was provided, ironically, by the people who least resembled the conception of an ideal citizen in classical democratic theory (Berelson, 1952). Ideally, it would be the people who are interested, attentive, and knowledgeable who should be changing, but by and large this was not true. The Berelson paradox, however, is more than irony. It also has practical implications for how political campaigns are conducted. Campaign managers and consultants probably had an intuitive awareness of which voters were susceptible to influence during a campaign, but the early voting studies made this explicit. Hence, it should not seem especially surprising if political campaigns are geared and conducted at a relatively low level (Quirk, 1989).

Although the Berelson paradox was firmly grounded in data, we want to reopen the matter by examining not only data from more recent U.S. election studies, but also by comparing U.S. results to those from Sweden. The choice of the United States and Sweden is not haphazard. Among the western democracies, Sweden and the United States are as different as any two systems. Sweden has a stable, multiparty system with representation in the unicameral *Riksdag* (Parliament) based on proportional representation. By contrast, the United States has two weak, undisciplined parties, representation in Congress based on the "winner take all" principle in separate constituencies, and a strong presidency. While we did not doubt that the Berelson paradox would be replicated in the United States, the process might indeed work differently in a system such as Sweden's (Holmberg, 1981; Petersson, 1977). In this article, we shall focus our analysis on within-campaign changers with only cursory attention given to across-election volatility. However, we have analyzed across-election volatility in these two countries and are confident that our observations generalize to that situation.

Data Sources

Our analyses are based on ten election studies done in the United States (1952–88) and seven in Sweden (1968–88). In each of the U.S. studies, conducted by the Survey Research Center/Center for Political Studies of the University of Michigan, a representative sample of adults was interviewed during the eight-week campaign period before the election and, whenever possible, after the election. The intention question in the preelection interview asks people directly, "Who do you think you will vote for in the election for president?" After the election, the same individuals are asked who they voted for in the presidential election.

The design for the Swedish election studies is somewhat different. A large sample is drawn and half the people are randomly chosen to be given a lengthy interview during the four-week campaign period before the election and the other half a similar interview after the election. Those interviewed before the election are not reinterviewed after the election. Rather, they are sent a very brief mail survey, with a telephone follow-up if necessary, in which they are asked, among other questions, whether and how they voted. The intention and voting behavior questions in the Swedish surveys are direct and the same as those in the U.S. studies except that in Sweden the reference is to voting for a party rather than for a particular candidate. In Sweden, people vote by selecting a list of candidates for the Parliament who have been nominated by a particular party.

The number of people interviewed in the Swedish studies has averaged about 2,800. In this article, given the nature of our problem, our analyses are limited to those who were interviewed in the preelection period and then recontacted after the election. Our analyses are further focused on those people who expressed an intention to vote for a specific party when interviewed on the preelection survey. (The number of people included in the analysis is given in tables 2 and 6.)

The sample in each of the Swedish studies is a random sample drawn from the register of the Central Statistical Bureau. The target population consists of people 18 to 80 living in Sweden who are eligible to vote. The percentage of the sample who were actually interviewed has averaged about 80%, with a range from 74% (1976) to 88% (1968), for the seven Swedish election studies used in our analysis. The Survey Research Center of the Central Statistical Bureau draws the sample and conducts the interviews. The Election Studies Project at the Department of Political Science of the University of Göteborg is responsible for constructing the questionnaire and coding the responses. The election studies in Sweden have been financed by a series of grants from the Swedish government (for more information, see Alfredsson, 1987; Gilljam and Holmberg, 1990).

Overall Results

Tables 1 and 2 give the overall results relating voting intention and voting behavior in the United States and Sweden, respectively. Most people in both countries vote as implied by their intentions, and the degree of the link between intention and behavior is similar in the two countries.[1] In spite of the low percentage of volatility, some systematic variation occurred. For instance, in Sweden, only 2% of the strong partisans switched from intention to behavior, compared to 14% of those who did not identify with a party. The comparable figures for the United States are 4% and 8%.

Interest and Knowledge

We now turn to the central question about whether it is the people who are least interested and least knowledgeable who are shifting the most from intention to behavior, as was observed in the early U.S. election studies. As a matter of fact, it appears to work that way for the United States but not for Sweden. Table 3 gives the effect of interest on the intention-behavior relationship across ten U.S. and seven Swedish elections.[2]

---

[1] In interpreting table 2, it is essential to bear in mind the tradition of bloc politics in Sweden. The Communist and Social Democratic parties comprise the Socialist Bloc, while the Center, Liberal, and Conservative parties comprise the Bourgeois Bloc. As can be seen in table 2, a majority of the intention-behavior changes are to another party in the same bloc. Switching from one bloc to the other is relatively rare and more infrequent than switching from one party to the other (Holmberg, 1984; Granberg and Holmberg, 1988).

[2] The Swedish question on interest in politics provides four alternatives and the U.S. question on interest in the political campaign uses three alternatives. Given the marginal distributions, there did not seem to be any fair way of collapsing categories so they

The results for interest as an intervening variable in the United States are clear-cut. Those who said they were "very much interested" in the political campaign were least likely to change from intention to behavior (average across years = 4%), followed by those who were "somewhat interested" (7%), and those who were "not much interested" (10%). Because of the consistency across years, one can have considerable confidence that the effect shown for the United States is real and dependable.

On the other hand, the effect of interest as a moderating variable in the intention-behavior relationship for Sweden is weak and not consistent across years. In 1985, for instance, of those who said they had a very high interest in politics, 94% had the same intention and behavior, compared to 92% of those with a quite high interest in politics, 93% for those not especially interested, and 96% for those not at all interested in politics. Combining across seven elections, the percentage whose intention and behavior were different varied within a very narrow range for the four interest levels, as can be seen in table 3. Taking the results for the two countries together, we observe a system by problem interaction, that is, the findings look noticeably different within the two political systems. The results for Sweden indicate that there is nothing inevitable or inexorable about the inverse relationship between involvement and volatility that has long been thought to exist in the United States (Berelson, Lazarsfeld, and McPhee, 1954).[3]

When we consider knowledge rather than interest as an intervening variable in the intention-behavior link, the results are similar. The 1985 Swedish study and the 1960 U.S. study contain several indicators of how well informed people are. The U.S. questions asked for knowledge about the two presidential candidates, while the Swedish survey had five knowledge questions about disputed issues and four questions dealing with the party affiliation of high-ranking party members. When the items were combined to form scales, the resulting knowledge indexes had reliability coefficients of .79 for the United States and .77 for Sweden (Cronbach's alpha).

Our concern is whether knowledge appears to moderate the intention-behavior relationship. That is, does the propensity to change from intention to behavior vary as a function of one's level of political

---

would be comparable for the two countries. A much larger percentage of people in the United States chose the highest interest alternative (of three) than people in Sweden (choosing among four alternatives). Thus, the highest interest group in Sweden is probably closer to the maximal end of a hypothetical interest continuum.

[3] The interest question in the U.S. studies pertains to interest in the campaign, while the interest item in the Swedish surveys deals with a more general interest in politics. That *could* be the source of the difference between the United States and Sweden in table 3, but we do not think so.

Table 1. The Intention-Behavior Relationship in U.S. Presidental Elections, 1952–88

	Intended Vote/Actual Vote							
	Dem/Dem %	Dem/Rep %	Rep/Dem %	Rep/Rep %	Total	N	r	Percentage Consistent
1988	42	4	4	50	100	1,090	.85	92
1984	39	2	2	57	100	1,269	.91	96
1980	43	5	1	51	100	719	.87	93
1976	47	3	3	47	100	1,355	.87	94
1972	28	2	4	66	100	1,384	.85	93
1968	41	3	4	52	100	763	.86	93
1964	68	4	1	27	100	987	.88	95
1960	43	2	4	51	100	1,227	.89	94
1956	39	4	2	55	100	1,130	.86	93
1952	39	3	3	55	100	1,039	.88	94
Average	43	3	3	51	100	1,096	.87	94

NOTE: The calculations in this table are based only on those people who expressed an intention, on the preelection survey, to vote for one of the two leading candidates, *and* who reported, on the postelection survey, having voted for one of the two leading candidates. Because of rounding, the percentage in the right-hand column does not always equal exactly the sum of columns 1 and 4.

Table 2. Subsequent Vote of People Intending to Vote for One of the Parties in Sweden's Parliament

Voting Intention		Percentage Voting For								
		Communist	Social Democrat	Greens	Center	Liberal	Conservative	Other	Total	*N*
Communist	1988	96	4	0	0	0	0	0	100	27
	1985	93	2	5	0	0	0	0	100	44
	1982	94	3	0	3	0	0	0	100	33
	1979	94	3	—	0	3	0	0	100	34
	1976	84	13	—	3	0	0	0	100	31
	1973	91	9	—	0	0	0	0	100	34
	1968	91	9	—	0	0	0	0	100	11
Social Democrat	1988	2	97	1	0	0	0	0	100	352
	1985	2	97	0	0	1	0	0	100	423
	1982	2	97	0	0	0	1	0	100	481
	1979	1	97	—	0	1	1	0	100	441
	1976	1	96	—	2	1	0	0	100	461
	1973	1	96	—	3	0	0	0	100	404
	1968	1	98	—	1	0	0	0	100	596
Center	1988	0	3	1	93	1	1	1	100	80
	1985	0	1	0	88	3	1	7	100	100
	1982	0	0	0	95	0	4	1	100	117
	1979	0	4	—	90	1	4	1	100	146

	1976	1	4	—	90	2	2	1	100	172
	1973	0	3	—	94	0	3	0	100	186
	1968	0	2	—	94	2	2	0	100	158
Liberal	1988	0	2	0	3	88	6	1	100	101
	1985	0	4	0	3	92	1	0	100	76
	1982	0	4	0	4	83	7	2	100	46
	1979	1	5	—	2	84	8	0	100	101
	1976	0	0	—	5	89	6	0	100	98
	1973	0	4	—	4	87	5	0	100	74
	1968	0	2	—	7	88	2	1	100	154
Conservative	1988	0	1	0	2	2	95	0	100	121
	1985	0	1	0	1	7	91	0	100	209
	1982	0	1	1	1	2	95	0	100	214
	1979	0	0	—	1	4	95	0	100	158
	1976	0	2	—	2	2	94	0	100	122
	1973	0	1	—	2	1	96	0	100	107
	1968	0	2	—	4	3	91	0	100	111
Greens	1988	3	6	81	3	5	0	2	100	63

NOTE: The percentages in this table are row percentages. The people included in this analysis were those who intended to vote for one of the seven largest parties, who later reported voting for one of these parties, and who were validated voters. The Environmental Party first became a factor in Swedish politics in 1982 and did not make it above the 4% threshold, required for gaining seats in the Paraliament, until 1988. Intender groups shown only consist of people intending to vote for one of the parties in Parliament. The seventh largest party in Sweden in recent years has been the Christian Democratic party, which has not yet succeeded in reaching the 4% threshold. Combining across these seven election studies, there were 124 people who intended to vote for the Christian Democratic party and, of these people, 105 (85%) later reported having done so.

Table 3. Percentage of Intention-Behavior Switching as a Function of Interest in Seven Swedish and Ten U.S. Election Studies

	Level of Expressed Interest in Politics			
Sweden	Not At All	Not Especially	Quite High	Very High
1988	3	9	6	1
1985	4	7	8	6
1982	3	6	6	4
1979	6	6	8	4
1976	8	8	7	4
1973	11	6	5	5
1968	5	5	4	6
Average	6	7	6	4
N	390	2,319	2,756	771

	Level of Expressed Interest in Election Campaign		
United States	Not Much	Somewhat	Very Much
1988	10	9	4
1984	8	5	1
1980	11	9	2
1976	10	8	4
1972	11	7	6
1968	14	7	5
1964	7	5	5
1960	13	6	4
1956	11	8	4
1952	7	8	3
Average	10	7	4
N	1,734	4,648	4,708

NOTE: As in tables 1 and 2, the percentages in this table are based only on those people who expressed an intention and later reported their vote. Each percentage shown in the table is the percentage of the people at a given level of interest who changed from intention to behavior.

knowledge? As with interest, the answer seems to vary as a function of the political system. In Sweden, there is virtually no relationship between level of knowledge and propensity to change from intention to behavior. The percentage of changers varies only between 4% and 7% for the five knowledge levels, with the most intention-behavior changing occurring among the highest and the moderately low levels of knowledge. So the finding for Sweden is a flat function with only minor deviations. In Sweden, the least knowledgeable citizens are clearly not the most volatile.

The results for the United States are similar, but with one major exception or outlier. The people in the United States with the least knowledge were unusually likely to be intention-behavior changers (20%). The level of intention-behavior changing was low with little variation among the next three levels of knowledge (5%, 6%, and 5%), and those with the most knowledge were the least likely to change from intention to behavior (3%). These results can also be seen as a system by problem interaction effect in which the relationship between level of knowledge and intention-behavior volatility is noticeably different for Sweden and the United States.

Because the measures of knowledge were quite different and taken at different times in Sweden and the United States, we also considered an alternative measure of political knowledge that has been used in six, more recent U.S. election studies (1968–1988). After the interview, the interviewer is asked to rate the respondent's "general level of information about politics and public affairs" using a five-point rating scale.[4]

We compared the percentage of intention-behavior changers among people who were rated above average, average, and below average by this interviewer-supplied measure of political knowledge. As shown in table 4, those rated above average in knowledge in each study were less likely to be intention-behavior changers than those rated below average. Combining across six elections, of those rated below average in knowledge, 11% were intention-behavior changers, compared to 7% of those who were rated average, and 5% of those rated above average in political knowledge. Thus, by another measure of political knowledge, we see that, in the United States, those with little political knowledge are more likely to change from intention to behavior than people with more knowledge. Yet, there is no reason to think it works that way in Sweden.

That the results for interest and knowledge would appear similar is not surprising in that these two moderating variables are not independent of one another. In each survey, interest and knowledge are positively correlated, + .44 in Sweden in 1985 between political interest and the knowledge index, + .47 in the United States in 1960 between interest in the campaign and the knowledge index, and + .44 in the United States in 1984 between interest in the campaign and inter-

---

[4] This is generally done after the preelection interview. When this was done after both the pre- and postelection interviews in 1968, the two ratings correlated at + .75 (N = 1,319). As one would expect, that correlation was somewhat higher if the ratings were made by the same interviewer ($r = + .77$) than if by different interviewers ($r = + .61$).

Table 4. Percentage of Intention-Behavior Switching as a Function of Interviewer-Rated Knowedge in Six U.S. Election Studies

Political Knowledge	1968	1972	1976	1980	1984	1988	Average	N
Above average	5	6	4	3	3	7	5	3,069
Average	5	5	9	8	5	8	7	2,418
Below average	14	13	10	13	5	9	11	1,132

NOTE: As in table 1, the percentages in this table are based only on those people who expressed an intention and later reported their vote. Each percentage in the table is the percentage of the people at a given level of knowledge who changed from intention to behavior.

viewer-rated knowledge. Therefore, we wanted to see what happens when interest and knowledge are given equal weight and combined into an index. This was done for each of these three surveys, and then the distribution was divided as nearly as possible into a 30-40-30 split on the interest-information index.

The results in table 5 reinforce earlier findings. Within the United States, those low in interest and knowledge are overrepresented among the intention-behavior switchers, while in Sweden, those high in interest and knowledge are overrepresented among the intention-behavior switchers.

An Interest-Partisan Typology

An alternative approach in the search for moderating variables in the intention-behavior relationship focuses on the intersection of interest and partisanship. In a methodological discussion of property space, Allen Barton (1955) suggested a typology considering political interest and party identification simultaneously. Those who identified with a party and had considerable interest in politics were called *partisans*. People with high interest but who did not identify with a party were called *independents*. Those who identified with a party but lacked interest in politics were called *habituals* and, finally, nonidentifiers who lacked interest were called *apathetics*. Such labels may be descriptive, but whether their implications are accurate (e.g., whether the habituals really are habitual voters) is an empirical question. This typology seemed potentially useful for our purposes.

A similar typology was devised and used by Olof Petersson (1977) in his analysis of the 1976 Swedish election. He described four types: *partiaktiva* (interested partisans), *fribytare* (interested nonpartisans),

Table 5. Intention-Behavior Changing in Three Election Studies and Position on an Interest-Knowledge Index

Study	Interest-Knowledge Index	Percentage of Intention-Behavior Changers (1)	Percentage of Voters (2)	Difference (3)
Sweden 1985	High	37	27	+10
	Medium	35	44	−9
	Low	28	29	−1
	Total	100	100	
U.S. 1984	High	7	24	−17
	Medium	61	54	+7
	Low	32	22	+10
	Total	100	100	
U.S. 1960	High	19	30	−11
	Medium	32	42	−10
	Low	49	28	+21
	Total	100	100	

NOTE: Column 1 gives the percentage of the intention-behavior changers who are at a given level on the interest-knowledge index. Column 2 gives the percentage of all those who expressed an intention and reported their vote. Column 3 gives the difference between the two percentages.

vaneröstare (uninterested partisans), and *passiva* (uninterested nonpartisans), and showed that differentiating people in this way could be useful in political analysis. For instance, the four categories differed noticeably in how well their vote could be predicted from class and ideology, how much political trust they had, and how likely they were to split their vote at the local and national levels. Most suggestive for our purposes, Petersson showed the four types differed significantly in volatility. Curiously, the interested nonpartisans (*fribytare*) were the most volatile across elections and also the most likely to report having decided relatively late during the election campaign (Petersson, 1977; Esaiasson, 1985).

Tables 6 and 7 show the trend in the percentage of these four types across time in Sweden and the United States. In Sweden, things are quite stable between 1968 and 1988 with perhaps a slight increase in the proportion of interested nonpartisans and uninterested nonpartisans, along with a corresponding decrease in the percentage of interested

Table 6. Four Types of Swedish Voters (in percent)

	Political Interest and Partisanship					
	High Interest Partisans (*Partiaktiva*)	Low Interest Partisans (*Vaneröstare*)	High Interest Nonpartisans (*Fribytare*)	Low Interest Nonpartisans (*Passiva*)	Total	*N*
1988	30	20	21	29	100	1,172
1985	33	21	18	28	100	1,228
1982	36	23	17	24	100	1,269
1979	36	24	19	21	100	1,170
1976	36	25	17	22	100	1,199
1973	35	24	16	25	100	1,158
1968	36	28	14	22	100	1,345

Note: The table percentages are row percentages giving the percentage in each year for each of the four types, based on the political interest and party identification questions.

Table 7. Four Types of U.S. Voters (in percent)

	Campaign Interest and Partisanship					
	High Interest Partisans (Partisans)	Low Interest Partisans (Habituals)	High Interest Nonpartisans (Independents)	Low Interest Nonpartisans (Apathetics)	Total	*N*
1988	20	44	8	28	100	1,995
1984	21	43	8	28	100	2,229
1980	20	43	10	27	100	1,546
1976	27	36	10	27	100	2,862
1972	22	42	9	27	100	2,696
1968	27	42	12	19	100	1,544
1964	32	44	7	17	100	1,553
1960	30	45	8	17	100	1,902
1956	24	49	6	21	100	1,748
1952	28	46	9	17	100	1,762

NOTE: The table percentages are row percentages giving the percentage in each year for each of the four types, based on the campaign interest and party identification questions.

Table 8. Percentage of Five-Party Intention-Behavior Changers in Sweden

	Political Interest and Partisanship			
	High Interest Partisans (*Partiaktiva*)	Low Interest Partisans (*Vaneröstare*)	High Interest Nonpartisans (*Fribytare*)	Low Interest Nonpartisans (*Passiva*)
1988	2	8	12	10
1985	6	4	11	9
1982	4	1	10	15
1979	5	4	14	8
1976	4	6	15	13
1973	3	5	12	11
1968	2	3	12	11
Combined	4	4	12	11
N	2,692	1,731	832	961

NOTE: The table gives the percentage of voters of each type who switched from intention to behavior in each year, based only on those who stated an intention to vote for one of the five main parties on the preelection interview and who reported voting for one of those parties on the postelection survey. Thus, the *N*s of table 8 are smaller than the *N*s of table 6.

and uninterested partisans (Gilljam and Holmberg, 1990). In the United States, there is a rather sudden increase in the percentage of apathetics between 1968 and 1972. The percentage of interested partisans seems to have dropped across years, especially if one compares 1964 and 1988 (Wattenberg, 1986 and 1987).

Tables 8 and 9 show how the percentage of intention-behavior changers varies across the four types. In Sweden, the uninterested partisans and the interested partisans have been about equally unlikely to change from intention to behavior. Interested nonpartisans and uninterested nonpartisans were considerably more likely to be intention-behavior changers and did not differ significantly from each other.

The results for the United States (table 9) present a rather different picture. The apathetic or passive, uninterested nonpartisans are the most likely to be intention-behavior changers followed by the uninterested partisans, the interested nonpartisans, and the interested partisans. Thus, with regard to intention-behavior changing, different patterns emerge in Sweden and the United States. In Sweden, partisanship is the key, and interest is not important in determining the proba-

Table 9. Percentage of Two-Party Intention-Behavior Changers in the United States

	Campaign Interest and Partisanship			
	High Interest Partisans (Partisans)	Low Interest Partisans (Habituals)	High Interest Nonpartisans (Independents)	Low Interest Nonpartisans (Apathetics)
1988	4	8	7	13
1984	1	6	3	6
1980	3	9	1	13
1976	4	8	7	10
1972	6	8	7	8
1968	5	8	5	9
1964	4	6	6	5
1960	4	7	4	14
1956	2	9	10	9
1952	3	7	3	10
Combined	4	7	6	9
N	3,646	4,665	1,053	1,714

NOTE: The table gives the percentage of voters of each type who switched from intention to behavior in each year, based only on those who stated an intention to vote for one of the candidates nominated by the two main parties on the preelection interview and who reported voting for one of those candidates on the postelection survey. Thus, the *N*s of table 9 are smaller than the *N*s of table 7.

bility of intention-behavior changing. In the United States, on the other hand, interest appears to exert the stronger effect.[5]

Table 10 examines what percentage of the changers are of each type and how that compares with the percentage of that type among all voters. Among intention-behavior changers, the uninterested partisans are clearly underrepresented in Sweden but overrepresented in the

---

[5] We also did a comparable analysis of cross-election volatility. In Sweden, the interested nonpartisans are the most likely (34%) to change across elections, followed by uninterested nonpartisans (30%). The uninterested partisans in Sweden are, once again, the least likely to change (8%), although their results are virtually the same as for the interested partisans (also 8%). Partisanship clearly has a strong effect on cross-election volatility in Sweden, while the effect of interest is quite weak. If anything, in Sweden those who are interested may be slightly more likely to change across elections than those who are uninterested in politics. In the United States, interested partisans are less likely (14%) to show cross-election volatility than uninterested partisans (20%). Among nonpartisans, both those lacking interest (27%) and those interested (27%) are quite likely to change.

Table 10. Comparison of the Percentage of Intention-Behavior Changers in Each of Four Types of Voters with the Percentage of Those Types of Voters in Sweden and the U.S.

	Sweden			United States		
	Percentage of Intention-Behavior Changers	Percentage of All Voters	Difference	Percentage of Intention-Behavior Changers	Percentage of All Voters	Difference
High interest partisans	27	43	−16	19	33	−14
Low interest partisans	19	28	−9	50	42	+8
High interest nonpartisans	27	13	+14	8	10	−2
Low interest nonpartisans	27	16	+11	23	15	+8
Total	100	100		100	100	
N	381	6,216		699	11,078	

NOTE: This table presents results combined across seven election studies (1968–88) for Sweden and across ten election studies (1952–88) for the United States. A negative number in the difference column means that the group is underrepresented among the intention-behavior changers, while a positive number in that column means the group is overrepresented among the intention-behavior changers.

United States. On the other hand, the interested nonpartisans are overrepresented among the intention-behavior changers in Sweden but underrepresented in the United States. Thus, the label given by Barton (1955) to the uninterested partisans, "habituals," seems to fit the behavior of uninterested partisans in Sweden quite well, but the comparable group in the United States less well.

Conclusions

It has long been recognized in normative democratic theory that people who are willing and ready to change between elections and within election campaigns provide flexibility and adaptability to a political system. Ideally, however, one would want people in this "floating bloc" to be interested, attentive, and informed as to what is going on in politics and specifically regarding the nature of the alternatives in electoral politics. It does not take us far beyond the data we have presented to suggest that the changers in Sweden more closely approximate that standard than do their counterparts in the United States.

The findings in our analyses of U.S. data constitute an extensive replication of the early findings pertaining to the Berelson paradox. That is, the flexibility and adaptability of the U.S. system depend upon citizens who do not fit the classical conception of the ideal citizen in a democracy. In the United States, intention-behavior changers tend to be disproportionately those who know little and lack an interest in politics.[6] Sweden, with its strong and more ideologically oriented party system, presents a very different picture. In Sweden, there is a category of people who identify with a party but lack an interest in politics (*vaneröstare*), and they show low volatility between elections and within election campaigns. The intention-behavior changers in Sweden, however, are *not* likely to be especially low in either political interest or political knowledge. In fact, the group with the highest proportion of intention-behavior changers in Sweden is the interested nonpartisans (*fribytare*).

[6] Because some research has shown a tendency for people to distort their reported vote in the direction of the winner (Granberg and Holmberg, 1986; Weir, 1975), it is worth considering whether this might be a confounding factor in our analysis. We detect no more than the slightest tendency in that direction. In the U.S. surveys (table 1), 95% of those who intended to vote for the candidate who won later reported voting for their intended candidate, compared to 93% of those who intended to vote for the candidate who lost. In the Swedish surveys (table 2), people intending to vote for the Social Democratic party in the years it lost (1976 and 1979) were about as likely to later report voting for their intended party as were people intending to vote for the Social Democratic party in the other years when it won. Generally, we do not see this as much of a problem when the time between voting and being asked in a survey how one voted is small. With the passage of time, however, memory dims and distortion becomes more likely.

The two political contexts of Sweden and the United States also differ dramatically in other ways. Sweden has a much higher rate of turnout, and its political parties are based more on social class (Inglehart, 1986; Korpi, 1981). Much of the support for Sweden's largest party, the Social Democratic party, has come from urban industrial unionized workers (Holmberg, 1981). Many of them identify with the Social Democratic party, are not highly involved in politics, and show very low volatility. The party system in Sweden has changed little across numerous elections and has been one of the most stable in Europe.[7]

In contrast, when low involvement voters in the United States indicate an intention to vote for a particular candidate, they may know relatively little about the opposing candidates. They may lack an abiding partisan commitment and their answer to the intention question may be somewhat more flippant than it is reflective of a firm decision. If so, it follows that they subsequently show a relatively high level of volatility.

In the United States, the intention-behavior switchers tend to be those who are apathetic and poorly informed.[8] This obviously is relevant when judging how a political system is working relative to normative democratic theory. At the practical level, when candidates and campaign managers are aware of this tendency, it obviously can have an effect on how a political campaign is geared and at what level of sophistication it is aimed. This is closely linked to what has recently been called the "paradox of mass politics" (Neuman, 1986).

The question remaining is why the process appears to work so differently in Sweden. Part of the answer seems to lie in the strong party system in Sweden where the principal actors are the political parties rather than the individual candidates nominated by the parties. Serious democratic theorists, such as V. O. Key (1949), Giovanni Sartori (1976), and Walter Dean Burnham (1989), have stressed the vital importance of the party and have argued forcefully that "pure candidate-domination of elections is a recipe for irresponsible, unaccountable performance in office and . . . for domination of the political order by the haves at the expense of the have-nots'' (Burnham, 1986:xiii). Based on our analyses, we can add that candidate domination of elections may also be a recipe for the Berelson paradox in which the more ignorant and apathetic citizens are most likely to be influenced during an election campaign, while this is not true in a system in which the political parties are the principal actors.

---

[7] In the most recent Swedish elections, the party system appears to be loosening up somewhat. According to a variety of indicators, the traditional "grip" of the five old parties on the voters is not as tight as it once was (Gilljam and Holmberg, 1990; Holmberg, 1984; Holmberg and Gilljam, 1987). One sign of this is the fact that in the 1988 election, the Environmental Party, the Greens, exceeded the 4% threshold, thereby gaining 20 seats in the Parliament and becoming the first new party in the Parliament in over 70 years.

[8] Kelley (1983), in his cleverly titled chap. 8, "Rule by the Worst of the Many," has addressed the same questions as ours, albeit with a different method. His results are compatible with ours, at least in certain respects. He observed that swing voters in the U.S. elections of 1964 and 1972 tended to be poorly educated, uninterested in politics, politically inactive, inattentive, and uninformed. However, Kelley also doubts that there were enough of these uninformed and apathetic switchers, especially in these landslide elections, for their votes to have been decisive.

References

Alfredsson, Iris (1987)
''The Swedish election studies: Studies of 30 years of Swedish electoral behavior.'' Swedish Social Science Data Service, University of Göteborg, 17–27.
Barton, Allen (1955)
''The concept of property space in social research.'' Pp. 40–53 in P. Lazarsfeld and M. Rosenberg (eds.), The Language of Social Research. New York: Free Press.
Berelson, Bernard (1952)
''Democratic theory and public opinion.'' Public Opinion Quarterly 16:313–30.
Berelson, Bernard, Paul Lazarsfeld, and William McPhee (1954)
Voting: A Study of Opinion Formation in a Presidential Campaign. Chicago: University of Chicago Press.
Burnham, Walter Dean (1986)
''Foreword.'' Pp. ix–xiv in M. Wattenberg, The Decline of American Political Parties 1952–1984. Cambridge, MA: Harvard University Press.
——— (1989)
''The Reagan heritage.'' Pp. 1–32 in G. Pomper (ed.), The Election of 1988: Reports and Interpretations. Chatham, NJ: Chatham House.
Campbell, Angus, Philip Converse, Warren Miller, and Donald Stokes (1960)
The American Voter. New York: Wiley.
Esaiasson, Peter (1985)
''Partiledareffekter även i Sverige?—Partiledarpopularitetens betydelse för valresultatet: Sverige ur ett jämförande perspektiv.'' Statsvetenskaplig Tidskrift 2:105–21.
Gilljam, Mikael, and Sören Holmberg (1990)
Rött Blått Grönt: En Bok om 1988 Års Val. Stockholm: Bonniers.
Granberg, Donald, and Sören Holmberg (1986)
''Prior behavior, recalled behavior, and the prediction of subsequent voting behavior in Sweden and the U.S.'' Human Relations 39:135–48.
——— (1988)
The Political System Matters: Social Psychology and Voting Behavior in Sweden and the United States. Cambridge: Cambridge University Press.
Holmberg, Sören (1981)
Svenska Väljare. Stockholm: Liber.
——— (1984)
Väljare i Förändring. Stockholm: Liber.
Holmberg, Sören, and Gilljam, Mikael (1987)
Väljare och Val i Sverige. Stockholm: Bonniers.
Inglehart, Ronald (1986)
''Intergenerational changes in politics and culture: The shift from materialist to postmaterialist value priorities.'' Pp. 81–105 in R. Braungart and M. Braungart (eds.), Research in Political Sociology, vol. 2. Greenwich, CT: JAI Press.

Kelley, Stanley (1983)
Interpreting Elections. Princeton, NJ: Princeton University Press.
Key, V. O. (1949)
Southern Politics in State and Nation. New York: Knopf.
Korpi, Walter (1981)
Den Demokratiska Klasskampen: Svensk Politik i Jämförande Perspektiv. Stockholm: Tiden.
Lazarsfeld, Paul, Paul Berelson, and Hazel Gaudet (1944)
The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign. New York: Columbia University Press.
Neuman, W. Russell (1986)
The Paradox of Mass Politics: Knowledge and Opinion in the American Electorate. Cambridge, MA: Harvard University Press.
Petersson, Olof (1977)
Valundersökningar-Rapport 2: Väljarna Och Valet 1976. Stockholm: SCB och Liber.
Quirk, Paul J. (1989)
"The election." Pp. 63–92 in M. Nelson (ed.), The Elections of 1988. Washington, DC: Congressional Quarterly.
Sartori, Giovanni (1976)
Parties and Party Systems: A Framework for Analysis. Cambridge: Cambridge University Press.
van der Eijk, Cees, and Broer Niemöller (1983)
Electoral Change in the Netherlands: Empirical Results and Methods of Measurement. Amsterdam: CT Press.
Wattenberg, Martin (1986)
The Decline of American Political Parties 1952–1984. Cambridge, MA: Harvard University Press.
——— (1987)
"The hollow realignment: Partisan change in a candidate-centered era." Public Opinion Quarterly 51:58–74.
Weir, Blair (1975)
"The distortion of voter recall." American Journal of Political Science 19:53–62.

CONTEXT EFFECTS FOR CENSUS MEASURES OF RACE AND HISPANIC ORIGIN

ELIZABETH MARTIN,
THERESA J. DEMAIO, and
PAMELA C. CAMPANELLI

Abstract This paper reports on the results of a split-ballot experiment conducted in 1987 to test alternative versions of the decennial census long form. Two forms were randomly assigned and self-administered in group sessions involving a total of 515 respondents. The order of race and Hispanic origin items was experimentally manipulated. The standard long form asks race, then Hispanic origin. The experimental form reversed the order of the items in order to reduce perceived redundancy, and to create a more restricted frame of reference for the race item. The objectives of the context manipulation were (1) to reduce item nonresponse for the Hispanic origin item, and (2) to reduce reporting of "Other race" by Hispanics in the race item. Objective (1) was met. Objective (2) was met for Hispanics born in a U.S. state, but not for immigrants. The results are interpreted as reflecting a process of acculturation that affects how Hispanic respondents apply U.S. racial categories "white" and "black" in the census.

The survey methodology literature contains much evidence of the effect of question sequencing on response to questions. This work has, in large part, involved attitude items (see Schuman and Presser, 1981; Turner and Martin, 1984). The effect of question order may derive from the context invoked by prior questions, which may influence respondents' frame of reference or suggest different interpretations of the question.

ELIZABETH MARTIN is Chief, Center for Survey Methods Research, Bureau of the Census. THERESA J. DEMAIO is Social Science Statistician and PAMELA C. CAMPANELLI is Survey Statistician, Center for Survey Methods Research, Bureau of the Census. This paper reports the results of research undertaken by Census Bureau staff. The views expressed are those of the authors and do not necessarily reflect those of the Census Bureau. The authors gratefully acknowledge the statistical assistance of Robert Fay and helpful comments by Robert Johnson and Irwin Schreiner. Any errors are the authors'.

Public Opinion Quarterly 54:551–566 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-04/$2.50

Context effects can also operate on quasi-factual and factual items, but research in this area has been limited. The potential for this type of effect is especially pronounced when the concept being measured is somewhat unclear, and the respondent really is not sure what is being asked. In this paper, we explore the effect of question sequence on responses to two potentially ambiguous items on the decennial census questionnaire: race and Hispanic origin.

The Measurement of Race

Despite its familiarity, the concept of race is not a simple one. Racial classifications, both popular and scientific, are based on a mixture of principles and criteria, including national origin, tribal membership, religion, language, minority status, physical characteristics, and behavior. The criteria and categories for racial classification vary among cultures and over time. In the United States, we are accustomed to thinking in terms of two major races: black and white.

In this country, we tend to treat race as an objective, fixed characteristic of a person that is biologically inherited. This meaning of race is so ingrained that it may come as a surprise to learn that other cultures have very different conceptions of race (Harris, 1968; Marshall, 1968). For example, the racial categories recognized in Brazil are not the same as those used in the United States, even though its population also includes components with white European and black African origins. According to Harris (1968), numerous racial categories are combined to create hundreds of different racial terms, based primarily on physical characteristics. Racial descent is not the rule; full siblings whose appearances differ are of different races. In addition, race is not a fixed characteristic of a person, and it changes when a person achieves wealth, for example, which is one of the criteria for race.

Cultures in Central and South America use different racial categories and criteria from those used in the United States. This difference in racial classifications implies that "white" and "black" are not natural catgories of race for most Spanish-speaking people. This fact has implications for the consistency and meaningfulness of the answers Spanish-speaking immigrants may give to the census race question.

Even within the United States, there have been significant shifts in how Americans understand and categorize race, which have been reflected in changes in the race categories used in the census. Race was first measured as a separate item in the 1850 census, using the categories "white," "black," and "Mulatto," which were also used in 1860. In the 1870 and 1880 censuses, categories for "Chinese" and "Indian" were added, and in 1890, "Japanese" was added. The interpretation of race as a biological (yet observable) characteristic in these early censuses is clear. Enumerators were instructed in 1870 to code race through observation and to "[b]e particularly careful in reporting the class *Mulatto*. The word here is generic, and includes quadroons, octoroons, and all persons having any perceptible trace of African blood. Important scientific results depend upon the correct determination of this class. . . ." (U.S. Bureau of the Census, 1979:18). The 1890 census saw an elaboration of the "Mulatto" category, with separate categories for mulattoes, quadroons, and octoroons. In 1900, the attempt to distinguish mixed black and white categories was abandoned but taken up again in 1910, when "Mulatto" was again listed as a category. In 1930, "Mulatto" was dropped permanently, and "A person of mixed White and Negro blood was to be returned as Negro, no matter how small the percentage of Negro blood" (U.S. Bureau of the Census, 1979:52).

There were also changes in the census racial classification of Spanish-speaking persons, such as Mexicans. Enumerators in 1930 were instructed that the race of "all persons born in Mexico, or having parents born in Mexico, who were not definitely White, Negro, Indian, Chinese, or Japanese, would be returned as Mexicans" (U.S. Bureau of the Census, 1979:52). In 1940, this rule was changed, and Mexicans and other persons of "Latin descent" were to be classified as "white" unless they were definitely of some race other than white. The rule was changed again in 1980, and Hispanic entries, such as Puerto Rican and Mexican, were left in the "Other race" category.

Finally, there were additions to the Asian categories. In 1960, categories were added for Filipino, Hawaiian, Part Hawaiian, Aleut, and Eskimo (Alaska and Hawaii had become states in 1959). Eskimo and Aleut were dropped in 1970 but restored in 1980, by which time there were 9 separate Asian and Pacific Islander categories, including Asian Indian, Guamanian, Samoan, and Vietnamese.

Since we think of race as a stable, enduring characteristic, it is fascinating to find that no single set of racial categories has been used in more than two censuses, and most were only used once. The changes partly reflect real change in the composition of the U.S. population due to migration from Asia and Central and South America, and the expansion of U.S. territory to include new groups. Undoubtedly, political considerations and pressures played a part in the decisions to create and combine categories as well.

The fluctuations in census race categories also suggest the difficulty of devising a meaningful, objective classification of race. Marshall (1968) argues that all racial classifications are arbitrary and artificial to some extent, whether they are based on "scientific" or popular criteria. Problems arise when respondents do not share the race categories used by the Census Bureau. Evidence suggests this is the case for many Spanish-speaking persons.

In 1970, the Census Bureau began conducting the census by self-enumeration rather than personal visit enumeration. Personal enumeration was conducted only for households that did not mail back a form. The change in census-taking procedure meant that, after 1970, race was based mainly on self-identification rather than enumerator observation.[1]

In conjunction with the change in procedure, dramatic changes in the racial characteristics of the population were observed. The number of Hispanic origin persons classified as "Other race" rose from 700,000 in 1970 to 5.8 million in 1980 (U.S. Bureau of the Census, 1987b:100). The transition to self-enumeration contributed to this increase. In addition, in 1980 persons of Hispanic origin were no longer recoded as "white" if they reported themselves as "Other race," as had been done in the 1970 census.

Additional evidence links self-enumeration with higher "Other race" reporting by Hispanics. In 1980, 35% of Hispanics were identified as "Other race" in census questionnaires, but only 10% were reported as "Other race" in personal visit reinterviews conducted after the census (McKenney, Fernandez, and Masamura, 1985). A possible reason for the inconsistency is respondent confusion with the 1980 census race question (U.S. Bureau of the Census, 1987b:100), which did not explicitly ask for race. Many persons filled in the circle for "Other race" and wrote in a nationality from Central or South America. Another possible cause is interviewer behavior in the reinterview study. Race was to be self-reported by respondents using a flashcard listing the race categories. Possibly, interviewers changed "Other race" responses to "white" for respondents who "looked white." Or, possibly, Hispanic respondents gave different answers in a personal interview (quite probably conducted by a white interviewer) than they gave on the census questionnaire.

The Measurement of Hispanic Origin

The Hispanic origin item was included in the census for the first time in 1970. In 1970, the item appeared only on the long form (a more detailed and extensive set of questions that goes to one in six households), but in 1980 it was asked of everyone as a 100% item. Placement near the race item on the questionnaire in 1980 may have affected reporting for both items.

---

[1] Strictly speaking, it is not quite accurate to say that race is based on self-identification, since one respondent fills out the form for all household members. This means that for most of the population, race (and all other census information) is based on proxy reports. Nevertheless, it is the intent of the Census Bureau to measure each person's racial self-identification, and the 1990 form includes a printed instruction to "Fill ONE circle for the race that the person considers himself/herself to be."

Hispanic origin had the highest nonresponse rate of any 100% item in 1980. The computer allocation rate was 4.2% for short forms and 2.3% for long forms (U.S. Bureau of the Census, 1986:32). The item may have appeared redundant to respondents who had just reported "Other race" and written in a Hispanic nationality. Many non-Hispanic persons may have left the item blank because they did not understand it, or because they did not find a category that fit them.

Evidence from the 1970 and 1980 censuses suggests lack of understanding of the question. Siegel and Passel (1979) and Passel and Word (1987) document extensive misreporting in 1970 in the "Central or South American" category by persons living in the central and southern United States. In 1980, fairly substantial numbers of people (particularly blacks in the South, Northeast, and Midwest) provided false positive reports of Mexican-American origin (U.S. Bureau of the Census, 1982). Respondents apparently wanted to indicate that they were American, and the Mexican-American response category was the only one that contained the word "American" (Passel and Word, 1987).

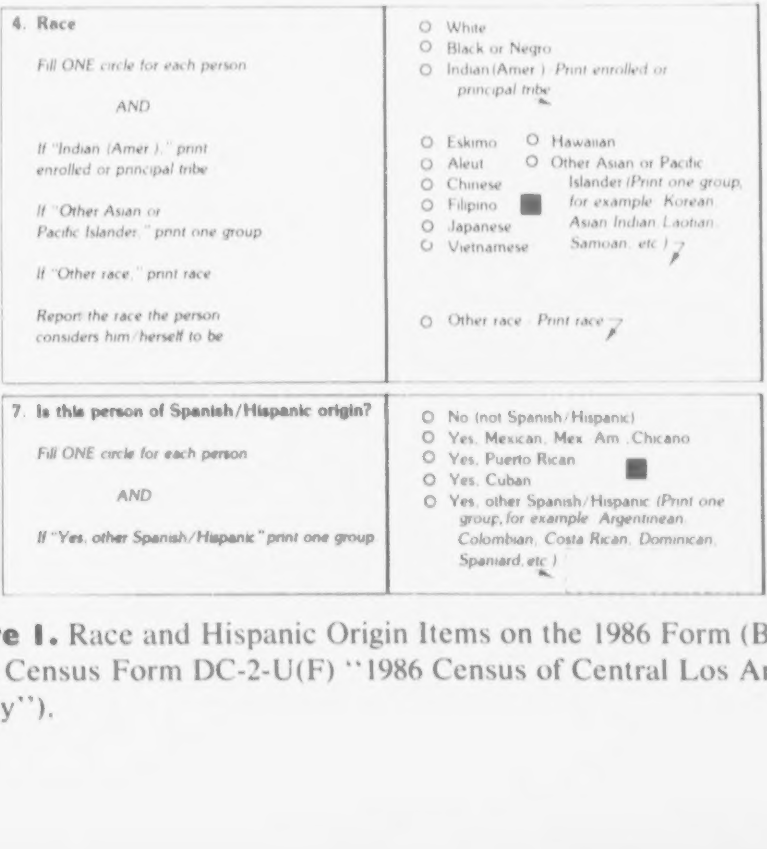
Some Hispanics were also confused by the question. A pattern of inconsistent reporting of "Other Hispanic" origin was observed between the 1980 census and the Content Reinterview Study (U.S. Bureau of the Census, 1986; McKenney, Fernandez and Masamura, 1985). While reporting by Puerto Ricans, Mexicans, and Cubans was fairly consistent (that is, the census agreed with the reinterview for 85% or more of the persons in the reinterview sample), only 55% of persons reported as "Other Hispanic" in the reinterview had been similarly classified in the census; the others had been reported mostly as Mexican (11%) and as non-Hispanic (31%).

Fernandez and Cresce (1986) found inconsistent reporting of Hispanic origin and Hispanic ancestry in the 1980 census. Many persons reported Hispanic origin but not Hispanic ancestry (e.g., 30% reported as Puerto Rican in origin did not report Puerto Rican ancestry). There were also over a half million reports of Hispanic ancestry in the census by persons who did not report Hispanic origin. The inconsistency may be evidence of the part-whole contrast effect described in the survey methodology literature (Schuman and Presser, 1981). This occurs when respondents "subtract out" their response to a previous question when answering a subsequent, more general one. Respondents to the ancestry question had already reported their Hispanic origin; when faced with a very similar question, they may have reinterpreted it to be nonredundant: "Besides what you already told me about, what is your ancestry?" Thus, these seemingly inconsistent reports may be evidence of respondents' attempts to make sense of the questions and provide nonredundant information.

Additional support for this interpretation is provided by an experiment that varied the contiguity of the race and Hispanic origin items (see McKenney, Cresce, and Johnson, 1988; U.S. Bureau of the Census, 1987a). The Hispanic origin item nonresponse rates ranged from 20% to 30% when placed immediately after race, compared to 13% to 17% when placed three items after race. This suggests that, with consecutive sequencing, the race item sets a frame of reference for the Hispanic origin item that accentuates the perceived redundancy of the latter item.

Hypotheses

We hypothesized that the high nonresponse for the Hispanic origin question and the reporting of "Other" in the race question may both result from context effects. The item that comes first may create a frame of reference that affects interpretation of the second item. The abundance of national origin groups listed as categories for race may encourage write-in entries of other nationality groups in that item. The Hispanic origin item then appears redundant, leading respondents to leave it blank (see fig. 1).



4. Race

Fill ONE circle for each person

AND

If "Indian (Amer.)," print enrolled or principal tribe

If "Other Asian or Pacific Islander," print one group

If "Other race," print race

Report the race the person considers him/herself to be

○ White
○ Black or Negro
○ Indian (Amer.) *Print enrolled or principal tribe*

○ Eskimo
○ Aleut
○ Chinese
○ Filipino
○ Japanese
○ Vietnamese
○ Hawaiian
○ Other Asian or Pacific Islander *(Print one group, for example: Korean, Asian Indian, Laotian, Samoan, etc.)*

○ Other race *Print race*

7. Is this person of Spanish/Hispanic origin?

Fill ONE circle for each person

AND

If "Yes, other Spanish/Hispanic" print one group

○ No (not Spanish/Hispanic)
○ Yes, Mexican, Mex. Am., Chicano
○ Yes, Puerto Rican
○ Yes, Cuban
○ Yes, other Spanish/Hispanic *(Print one group, for example: Argentinean, Colombian, Costa Rican, Dominican, Spaniard, etc.)*

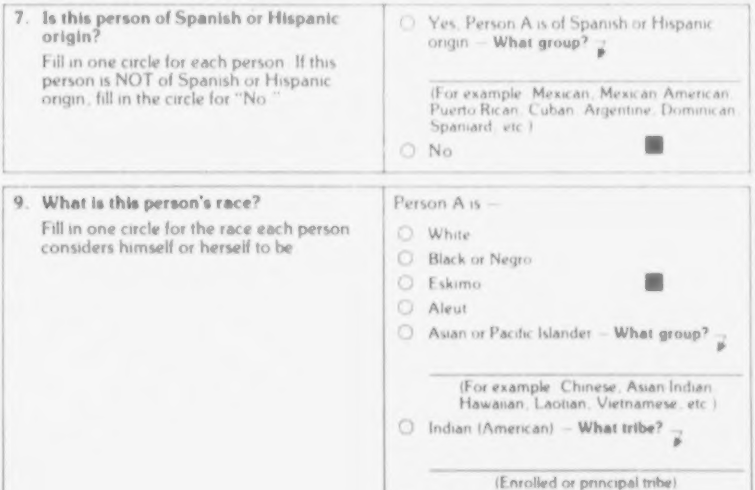
Figure 1. Race and Hispanic Origin Items on the 1986 Form (Bureau of the Census Form DC-2-U(F) "1986 Census of Central Los Angeles County").

We reasoned that the Hispanic origin item is less redundant if it precedes race. People who think their answer to the Hispanic origin question can be inferred from their response to the race question would be less likely to leave the Hispanic origin question blank if they answer it first. By reversing the order of the two items, we hoped to decrease item nonresponse for Hispanic origin (see fig. 2).

The reporting of "Other" in the race item may also be affected by the order of the two items. The majority of "Other" races which are written in are Hispanic nationalities. We thought that giving Hispanics a chance to report their Hispanic origin first would create a more restricted frame of reference for the race item, and reduce their reporting of "Other race." However, if Hispanics reject the categories "white" and "black," then their reporting of "Other" race may be insensitive to the order in which the items are asked (see DeMaio and Martin, 1987, for a statement of the hypotheses on which our revisions to the census form were based).

Methodology

This study was conducted as part of a program of research to improve the design of the decennial census long-form questionnaire. In this redesign effort, typographic and other layout changes were made to increase the consistency of the form's appearance, clearer instructions



7. **Is this person of Spanish or Hispanic origin?** Fill in one circle for each person. If this person is NOT of Spanish or Hispanic origin, fill in the circle for "No."	○ Yes, Person A is of Spanish or Hispanic origin — **What group?** (For example: Mexican, Mexican American, Puerto Rican, Cuban, Argentine, Dominican, Spaniard, etc.) ○ No
9. **What is this person's race?** Fill in one circle for the race each person considers himself or herself to be	Person A is — ○ White ○ Black or Negro ○ Eskimo ○ Aleut ○ Asian or Pacific Islander — **What group?** (For example: Chinese, Asian Indian, Hawaiian, Laotian, Vietnamese, etc.) ○ Indian (American) — **What tribe?** (Enrolled or principal tribe) ○ Other race — **What race?**

Figure 2. Hispanic Origin and Race Items on the Revised Form (Bureau of the Census Form DC-1409 "Census of the United States").

were provided on how to complete the census form, questions were reworded to simplify and clarify concepts, and questions were reordered to improve the flow and coherence of the census form (see DeMaio, Martin, and Sigman, 1987; Bates and DeMaio, 1989; for further discussion of the goals and methods of the research).

Figures 1 and 2 are facsimiles of the race and Hispanic origin items for the 1986 test census form, which was used as a control, and the revised form. As can be seen, several changes were made to these items, the principal one being that the items were reversed to test the hypotheses discussed above. In both forms, race and Hispanic origin were separated by other items (age in the revised form; age and marital status in the 1986 form). Other revisions were also made. Response categories for the Hispanic origin item were reordered, as they were throughout the revised form, so that "Yes" preceded "No." An instruction to fill in the "No" circle if the person was not Hispanic was added. There were also differences in response categories in the revised form, which reflected the thinking within the Census Bureau at the time regarding the content of these items in 1990. In the 1986 form, detailed Hispanic (for Hispanic origin) and Asian (for race) categories were listed separately. In the revised form, categories were combined, with an instruction to write in each person's specific group.

The revised and 1986 forms were compared in a series of split-panel experiments conducted in April, 1987, in about 30 group sessions organized by the Census Bureau Regional Offices in Boston, Dallas, Chicago, and Philadelphia. Participants in the sessions were volunteers recruited by Regional Office outreach staff through their contacts with community groups. Recruiting was focused on minority racial and ethnic groups, and people with relatively little education. Among the volunteers who appeared for the sessions, there were a total of 9 nonrespondents who did not attempt to fill out census forms (3 refusals, 2 persons with visual impairments, 2 with language problems, and 2 late arrivals). A total 515 people filled out forms for themselves and members of their households, providing data on 1,446 persons. Participants included people aged 18 to 80, members of different racial and ethnic groups, and people with various levels of education (see Parsley et al., 1989, for a detailed report of the results of this research).

Each session lasted an hour and a half, with 55 minutes allotted to fill out the census long form. (Forty-five minutes is the Census Bureau's official estimate of the time it takes an average household to fill out this form.) During each session, half of the participants were randomly assigned the 1986 form and the other half were assigned the revised form. Even though these respondents do not represent a probability sample, the randomization by form type permits us to make statistical comparisons between forms. We tested for statistically significant form differences using chi-square tests that take into account the clustering of persons within group sessions and within households (Fay, 1989). In the tables that follow, the chi-square values are Pearson chi-squares, calculated on the assumption of simple random sampling: the *J*s refer to jack-knife statistics that take into account clustering in the data and thus represent a more conservative test. Likelihood ratio chi-squares (identified as L^2 in the tables) were used to test the fit of alternative log-linear models.

Results

FORM DIFFERENCES IN DISTRIBUTIONS OF RACE AND HISPANIC ORIGIN

Despite revisions to the race question and response categories, race distributions are very similar for the two forms. As shown in the first section of table 1, each form identified about the same percentage of individuals in the categories of white, black, Asian or Pacific Islander, and other race. The revised form, however, identified a significantly larger percentage of American Indians. This result is due to sampling variability and is not a form effect.[2]

The second section of table 1 contains the distribution of responses to the Hispanic origin item, collapsed into the categories on the revised form for comparability. Again, the two forms obtained similar levels of reporting of Hispanic origin. Although based on small numbers, this result is encouraging because it suggests that the increased response to the item in the revised form (as reported below) did not alter the distribution.

ITEM NONRESPONSE

Results in table 2 support our hypothesis concerning item nonresponse to the Hispanic origin item. When Hispanic origin is asked first, item nonresponse drops to 9%, compared to 18% on the 1986 form. Nonresponse for the race item was not affected by the change in order, with a rate of 3% for the 1986 form and 4% for the revised form.

Because of the number of changes made to the form, we cannot be sure that the reversed order is the cause of the better response rate for Hispanic origin in the revised form. However, the pattern of missing data by race leads us to believe that the sequence of the items and

---

[2] Three American Indian respondents, each with large families, were randomly assigned to receive the revised from, whereas only one American Indian respondent, living alone, was assigned to receive the 1986 form.

Table 1. Distribution of Race and Hispanic Origin by Type of Form

	1986 Form	Revised Form
Race[a]		
White	53%	53%
Black	30	31
Asian or Pacific Islander	3	2
American Indian	—[b]	3
Other race	14	11
Total	100	100
Number of persons with complete data on race	684	715
Hispanic Origin[c]		
Hispanic	23%	19%
Non-Hispanic	77	81
Total	100	100
Number of persons with complete data on Hispanic origin	579	678

[a] $\chi^2 = 3.37$, *df* = 3, n.s., American Indian category excluded from test.
[b] Less than 1 percent.
[c] $\chi^2 = 2.49$, *df* = 1, n.s.

Table 2. Item Nonresponse Rates for the Race and Hispanic Origin Items by Type of Form

	1986 Form	Revised Form
Race[a]		
No response	3%	4%
Valid response	97	96
Total	100	100
N	703	743
Hispanic Origin[b]		
No response	18%	9%
Valid response	82	91
Total	100	100
N	703	743

[a] $\chi^2 = 1.31$, *df* = 1, n.s.
[b] $\chi^2 = 25.13$, *df* = 1, $p < .001$; $J = 2.78$, *df* = 1, $p < .01$.

Table 3. Item Nonresponse to Hispanic Origin Item by Race and Type of Form

	White	Black	Asian or Pacific Islander	Other Race
	1986 Form[a]			
No response on Hispanic origin	7%	36%	41%	4%
Valid response on Hispanic origin	93	64	59	96
Total	100	100	100	100
Number of persons with complete data on Race, excluding American Indians	361	204	22	96
	Revised Form[b]			
No response on Hispanic origin	8%	11%	0%	5%
Valid response on Hispanic origin	92	89	100	95
Total	100	100	100	100
Number of persons with complete data on Race, excluding American Indians	380	220	16	80

NOTE: Three-way interaction: $L^2 = 25.63$, $df = 2$, $p < .001$; $J = 1.93$, $df = 2$, $p < .05$.

[a] $\chi^2 = 98.40$, $df = 3$, $p < .001$; $J = 5.80$, $df = 3$, $p < .001$.

[b] $\chi^2 = 4.83$, $df = 3$, n.s.

the additional instruction to fill in the "No" circle were responsible. Results for the 1986 form (see the first section of table 3) shows that persons who reported their race as "black" or "Asian/Pacific Islander" were far more likely to leave the Hispanic origin item blank than persons reported as "white" or "other race." On the revised form, however, all race groups were equally likely to leave the Hispanic origin item blank (see second section of table 3).

A loglinear model confirms the presence of a three-way interaction (Nonresponse to Hispanic origin × Race × Type of form). This finding supports our hypothesis that the race item conditions respondents' understanding of the intent of the Hispanic origin item. Our interpretation is that blacks and Asians are especially likely to think their response ("No") can be inferred from their response to the race item, and hence they leave Hispanic origin blank. When Hispanic origin is first and there is an explicit instruction to fill in the "No" circle for

Table 4. Reporting of White versus Other Race Among Hispanics by Type of Form

	1986 Form	Revised Form
White	25%	39%
Other race	75	61
Total	100	100
Number of persons with complete data on Hispanic origin and race	122	111

NOTE: $\chi^2 = 4.76$, $df = 1$, $p < .05$; $J = .81$, $df = 1$, $p = .12$.

persons who are not Hispanic, persons of all races are likely to give a response.

CONTENT DIFFERENCES—RACE

Our second major hypothesis concerned the level of reporting of "Other race" by persons of Hispanic origin. For both questionnaire versions, the vast majority of write-in entries in "Other race" were Hispanic (over 90% for both forms). This finding is consistent with prior evidence.

The results suggest that the form of the questionnaire did affect Hispanic individuals' responses to the race item. As shown in table 4, persons who reported Hispanic origin were more likely to report their race as "white" in the revised form (39%) than in the 1986 form (25%) although this difference is not quite statistically significant with the more conservative test ($p = .12$). However, the extent of reporting "Other race" write-ins by Hispanics was substantial on both forms (61% for the revised form vs. 75% for the 1986 form). Thus, our goal of reducing reporting of "Other race" by Hispanics was only partially successful. These data suggest that most of the persons of Hispanic origin who participated in our test did not identify with any of the race categories listed in the census questionnaire.

Further analysis shows that the effect of context was restricted to Hispanics who were born in the United States. Table 5 presents responses to the race item, separately for people born in a U.S. state or outside the United States (i.e., in a foreign country or U.S. territory); the table includes data for persons of Hispanic origin only, and excludes those for whom information on place of birth is missing. The left section shows that race reporting for Hispanics who were not born in a U.S. state is unaffected by the version of the questionnaire. For

Table 5. Reporting of White versus Other Race Among Hispanics by U.S. Birth and Type of Form

	Born outside the U.S.[a]		Born in a U.S. State[b]	
	1986 Form	Revised Form	1986 Form	Revised Form
White	17%	18%	22%	74%
Other race	83	82	78	26
Total	100	100	100	100
Number of persons with complete data on place of birth	46	44	32	38

NOTE: Three-way interaction: $L^2 = 8.34$, $df = 1$, $p < .01$; $J = 1.04$, $df = 1$, $p = .08$.
[a] $\chi^2 = 0.01$, $df = 1$, n.s.
[b] $\chi^2 = 18.65$, $df = 1$, $p < .01$; $J = 2.21$, $df = 1$, $p = .01$.

this group, race is reported as "Other" and a Hispanic nationality is written in for over three-quarters of the cases on both forms. However, there is a very large context effect on race reporting for Hispanic persons who were born in a U.S. state. For U.S.-born Hispanics, race was significantly more likely to be reported as "white" in the revised form (74%) than in the 1986 form (22%). (The three-way interaction between form, place of birth, and race is marginally significant.)

This result suggests that most Hispanic immigrants simply do not apply the categories "white" or "black" to themselves, regardless of whether or not they have first reported their Hispanic origin. Race reporting for U.S.-born Hispanics, in contrast, appears more contextual. Results for this group are consistent with our initial hypothesis that giving Hispanic respondents a chance to report Hispanic origin before asking race would reduce the extent of "Other race" reporting.

These differences in race reporting for U.S.-born and immigrant Hispanics may reflect a process of acculturation. Our statistical results (namely, the interaction effect between place of birth, form type, and reported race for persons of Hispanic origin) imply that Hispanic respondents' assimilation and use of the U.S. race category "white" are complex, and depend both on context and on the category of person being described.[3] When Hispanic identity has already been acknowledged (in the revised form, by the prior question about Hispanic origin), Hispanics (especially those born in the United States) are more likely to report their race as "white." We should emphasize that we offer these substantive interpretations quite tentatively; further testing on larger and more representative samples is needed before they can be considered more than hypotheses.

---

[3] The reader may wonder how proxy vs. self-reports of race are affected by the form experiment. We do not have the data to examine this question directly. However, in order to shed light on possible proxy effects, we examined race reporting in 38 Hispanic households in which one or both parents and their children were living (nonrelatives

If they are replicated, our results imply that the measurement properties of census items about race and Hispanic origin are affected by context effects, and that the measurement properties of these items vary systematically over the population. This result is consistent with Johnson's (1990) finding that the measurement properties of different indicators of Hispanicity are not constant, but vary between first or second generation Hispanic immigrants versus others.

Discussion and Conclusions

Overall, these findings suggest that the revisions to the form tested in these experiments could result in improvements in data quality. We increased the response rate to the Hispanic origin item, which was one of the major hypotheses guiding this research on question sequence. This was done without affecting the substantive distribution, and without affecting the response rate for the race item.

One limitation of the current study is the number of experimental manipulations introduced in our revised questionnaire. Three experimental manipulations could affect the results reported here: the order of the Hispanic origin and race items, the added instruction to fill in the "No" circle for Hispanic origin, and the variations in response categories for race and Hispanic origin. A second limitation is the fact that results are based on small and purposively selected "samples." We are presently planning and conducting questionnaire experiments based on larger, randomly selected samples, including a questionnaire experiment planned as an integral part of the 1990 census.

Despite the limitations, our findings support some general conclusions about questionnaire design. The first is that redundancy is a problem from the respondents' point of view. Respondents apparently try to interpret questions so as not to provide redundant information. When questionnaire items really are partially or wholly redundant, a substantial amount of error and missing data can result.

A second conclusion is that our results call into question the naive

---

and other relatives in the household were ignored). In 92% of these cases, race was reported consistently for the entire nuclear family, which suggests there are not large proxy effects on race reporting.

assumptions about "error" with which we began our research. We aimed to improve data quality by, among other things, reducing reporting by persons of Hispanic origin in the "Other" category of the race item. As we analyzed our data we came to the point of view that the patterns of reporting reflect real differences in the interpretation of the meaning of race. We believe that the reporting of Hispanics in the "Other" race category is not an error, but a real perception by these respondents of their "correct" place in the racial classification structure—neither "white" nor "black." We agree with the point of view stated by Lieberson (1985:160):

> . . . some of the difficulties that researchers and census takers experience in using data on racial and ethnic groups are due not to problems of instrumentation or execution, pure and simple—such as might occur if the question was constructed incorrectly through, say, some vagueness or ambiguity of meaning. Rather some of the difficulties and inconsistencies reflect the processes of ethnic and racial change themselves; the "errors" are telling us something about the flux in the concepts and identifications themselves.

Our findings concerning context effects on race reporting by Hispanics may tell us something about the complex process of assimilation of U.S. racial categories. Our speculation is that Hispanic respondents may have their own "rules" for reporting race in the census for themselves and members of their households that do not conform to the Census Bureau's rules or the common understanding of race. The rule (roughly stated) might be: apply the U.S. racial category "white" to describe Hispanic persons with more "roots" in the United States, but only after acknowledging Hispanic identity. If there is such a rule, it has important implications for the meaning and the measurement of race in the census and in general, which we plan to explore in our future research on this subject.

References

Bates, Nancy A., and Theresa J. DeMaio (1989)
"Using cognitive research methods to improve the design of the decennial census form." Paper presented at the Census Bureau's Fifth Annual Research Conference.
DeMaio, Theresa J., and Elizabeth Martin (1987)
"Documentation of revisions to 1986 test census long form." U.S. Bureau of the Census, manuscript.
DeMaio, Theresa J., Elizabeth A. Martin, and Betsy Page Sigman (1987)
"Improving the design of the decennial census questionnaire." Paper presented at the annual meeting of the American Statistical Association, Survey Research Methods Section.
Fay, Robert (1989)
"CPLX: Contingency Table Analysis for Complex Sample Designs, Program Documentation." U.S. Bureau of the Census, manuscript.

Fernandez, Edward W., and Arthur R. Cresce (1986)
"Who are the other Spanish?" Paper submitted for presentation at the annual meeting of the Population Association of America, San Francisco.
Harris, Marvin (1968)
"Race." In International Encyclopedia of the Social Sciences 13:263–68.
Johnson, Robert A. (1990)
"Measurement of Hispanic ethnicity in the U.S. Census: An evaluation based on latent-class analysis." Journal of the American Statistical Association 85:58–65.
Lieberson, Stanley (1985)
"Unhyphenated whites in the United States." Ethnic and Racial Studies 8(1): 159–80.
Marshall, Gloria (1968)
"Racial classifications: Popular and scientific." Pp. 149–64 in M. Mead, T. Dobzhansky, E. Tobach, and R. Light (eds.), Science and the Concept of Race. New York: Columbia University Press.
McKenney, Nampeo R., Edward W. Fernandez, and Wilfred T. Masamura (1985)
"The quality of the race and Hispanic origin information reported in the 1980 Census." Proceedings of the Survey Research Methods Section (American Statistical Association), 46–50.
McKenney, Nampeo D., Arthur R. Cresce, and Patricia Johnson (1988)
"Development of race and ethnic items for the 1990 census." Paper presented at the annual meeting of the Population Association of America, New Orleans.
Parsley, Teresa L., Theresa J. DeMaio, Elizabeth Page Sigman, Elizabeth Martin, Pamela C. Campanelli, and Patricia Petrick (1989)
"Report on results of the classroom testing for the questionnaire design project." 1986 Test Census, Preliminary Research and Evaluation Memorandum no. 75. Washington, DC: U.S. Bureau of the Census.
Passel, Jeffrey, and David Word (1987)
"Problems in analyzing race and Hispanic origin data from the 1980 Census: Solutions based on constructing consistent populations from microlevel data." Paper presented at the annual meeting of the Population Association of America, Chicago.
Schuman, Howard, and Stanley Presser (1981)
Questions and Answers in Attitude Surveys. New York: Academic Press.
Siegel, Jacob S., and Jeffrey S. Passel (1979)
"Coverage of the Hispanic population of the United States in the 1970 Census: A methodological analysis." U.S. Bureau of the Census, Current Population Reports, series P-23, no. 82.
Turner, Charles F., and Elizabeth Martin (1984)
Surveying Subjective Phenomena. New York: Basic Books.
U.S. Bureau of the Census (1979)
Twenty Censuses: Population and Housing Questions, 1790–1980. Washington, DC: U.S. Government Printing Office.
——— (1982)
Preliminary Evaluation of Responses in the Mexican Origin Category of the Hispanic Origin Item. Persons of Hispanic Origin by State: 1980. 1980 Census of Population. Supplementary Report PC80-S1-7:14–17.
——— (1986)
Content Reinterview Study: Accuracy of Data for Selected Population and Housing Characteristics as Measured by Reinterview. 1980 Census of Population and Housing Evaluation and Research Reports, PHC80-E2.
——— (1987a)
"1986 national content test: Evaluation of Spanish/Hispanic question (Preliminary Draft)." Evaluation report prepared by the Ethnic and Hispanic Statistics Branch, Population Division.
——— (1987b)
1990 Planning Conference Series no. 12: Race and Ethnic Items. Washington, DC: U.S. Bureau of the Census.

ELECTORAL CONSEQUENCES OF "WINNING" TELEVISED CAMPAIGN DEBATES

PETER R. SCHROTT

Abstract This study explores the long-neglected impact of a candidate being perceived as the "winner" of political debates on individual voting behavior in West Germany. While there has been much research on the effects of debates on perceptions of candidate policy stands and personality attributes, there has been little work that attempts to gauge how much "winning" the debate helps a candidate at the polls. I investigate this process with straightforward probit models of the vote choice, including party identification and candidate evaluation, as well as whether the individual judged a particular candidate as the "winner," or "fared best" in the debate. These models are run for the 1972, 1976, 1980, and 1983 elections in West Germany. The results show that there is a consistent impact of "winning" the debate on individual voting choices in West Germany; controlling for party and other factors, there is a significantly higher probability of voting for the party of the politician the respondent feels won the contest. Debates, then, do ultimately make a substantial electoral difference.

Since the first televised presidential debate between Kennedy and Nixon in 1960, both the public and social scientists have become increasingly interested in following political campaign debates. For the public, televised debates are perhaps the only opportunity to watch and judge the candidates in a face-to-face contest. In these events the usual media "gatekeepers" (White, 1950) are bypassed, and the voters are given a chance to assess the issue positions, personal qualities, and programs of the candidates. For these reasons, debates typically draw large audiences, larger than any other campaign event (Kraus,

PETER R. SCHROTT is Assistant Professor of Communications at the Free University of Berlin/West Germany. The author thanks Steve Finkel, Lutz Erbring, and Dorothy Davidson for their comments and criticisms of this paper. Also appreciated is the helpful advice of Helmut Norpoth, Alan Abramowitz, and Shanto Iyengar at an earlier stage of this research. This article is a revised version of a paper delivered at the 1988 meeting of the Midwest Political Science Association.

Public Opinion Quarterly 54:567–585 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-04/$2.50

1962 and 1979; Graber, 1978). For the social scientist, debate research has centered on assessing the impact of these contests on the individual voter, on his or her evaluations and knowledge of the various participants, and on changes in vote intentions that may be due to information presented during the debate (Abramowitz, 1978; Bishop, Meadow, and Jackson-Beeck, 1978; Becker et al., 1980).

Numerous studies have analyzed televised debates and the electorate's reaction to them. We have learned, for example, that the audience is very interested in the event (Katz and Feldman, 1962), and gains some information about candidates' personalities and issue stands (Lang and Lang, 1962 and 1979; Miller and MacKuen, 1979; for the West German context see Baker, Norpoth, and Schönbach, 1981). Abramowitz (1978) studied the impact of the 1976 U.S. debates on the rationality of voting decisions and found strong evidence of a persuasion effect, in that the voters adopted the issue position taken by their preferred candidate. Many other studies also demonstrate information gains and reinforcement of preexisting partisan tendencies resulting from debate exposure (Weiss, 1976; O'Keefe and Mendelsohn, 1979). While it has been shown that people do learn from debates, and to a somewhat lesser extent, that debates affect the respondent's evaluation of the candidates (see also Kraus and Davis, 1976), it has never been conclusively demonstrated that debates affect vote choice. In this paper, I examine the direct and indirect influences that debates have had on West German voting behavior through the relatively unexplored concept of the debate "winner." The results show that by selecting a particular candidate as the winner, West German voters in all elections since 1972 were more likely to evaluate that politician more positively, more likely to evaluate his opponents more negatively, and ultimately more likely to vote for the winner's party. Debates, then, strongly influence vote choice in the West German context.

The Concept of the Winner

Most researchers argue that debates are single events among a multitude of campaign events and thus cannot be expected to have a strong effect on the electorate (Kraus, 1962 and 1979; Ranney, 1979), especially in lopsided elections when one party or candidate has an insurmountable lead. However, in a very close election even a slight debate effect might be enough to determine the outcome. For example, John F. Kennedy's victory in 1960 was attributed by some to his excellent debate performance (Katz and Feldman, 1962; Kraus, 1962 and 1979; Martel, 1983).

There might also be an indirect or intervening effect of winning due to the very nature of a debate. A debate is fundamentally a contest between two or more participants and is thus an integral part of the campaign "horse race" (Katz and Feldman, 1962), where it is possible for the public to see a "winner" and one or more "losers." In addition, debates usually take place close to election day and people recalling this event might not remember everything the participants said, but they might remember well who "won" and who "lost." It is this concept of the winner of the debate that might give the respective candidate a boost in the voter's perception of his overall evaluation. An undecided voter might base his or her decision on this last opportunity to get some information about the politicians and then simply vote for the perceived winner. Sears and Chaffee (1979) reviewed various studies on the 1976 presidential and vice presidential debates, and found a good bit of consensus with regard to the perceived winner. Almost all studies arrived at the same general outcome and, despite some marginal differences, the authors concluded that there was a net perception that Ford won the first debate and Carter the second.[1] Yet, this perception was attributed mainly to the media coverage of the debate, and less to the actual watching itself.

McLeod et al. (1979) found some impact of various debate variables on changes in presidential voter preference. The respondents in their sample who saw Carter as the debate winner became more involved in the campaign and showed increased affinity with the Democratic party. They also revealed vote preference shifts toward Carter. Yet the sample was rather small and made up mainly of young voters, and the authors did not control for party identification or candidate evaluation. Since other studies (for a review see Sears and Chaffee, 1979) found that people tended to perceive "their" candidate as the debate winner, the inclusion of party identification and candidate evaluation in a vote model seems a necessary condition to obtaining independent effects of winning a debate.

Most recent research into the concept of the winner has been done in the West German electoral context. Comparing the 1972 and 1976 debates, Baker, Norpoth, and Schönbach (1981) found a divergence between the political content of the debates and the recall of this content by the respondents. They also suggested that politicians with a more positive debate style were in a better position to win the debates. For the 1972 debate, Baker and Norpoth (1981) showed some information gain concerning the evaluation of the opposition parties. They also found a greater net improvement of his voting share among debate viewers for the perceived winner of the debates series, whereas the losers of the debates were found to gain among nonviewers. Still, they acknowledge that their data cannot tell whether there would have been a shift in voting behavior among debate viewers and nonviewers independent of the debate, and thus their study does not constitute clear-cut proof of the winner effect.

---

[1] In 1976 there were three presidential debates between Ford and Carter, and one vice presidential debate between Dole and Mondale.

The present investigation will take account of the various influencing factors, and I will show that winning the debates does improve not only the candidates' images, but also their chance of getting elected.

Data and Methods

Since 1972, a major feature of the electoral process in the Federal Republic of Germany is the televised debate between the leaders of the major parties represented in the Bundestag. It takes place three days before election Sunday, and the entire event is broadcast by the two national television networks, which preempt all other programming at that time. Since these events provide the public with a unique possibility to gather information on the competing parties, the party leaders, and their policy stands, many questions have been included in the German National Election Surveys to tap possible debate effects. For this analysis, panel data of the four election surveys in 1972, 1976, 1980,[2] and 1983[3] were used to estimate various models of debate impact.[4] The postelection waves were conducted about one to two weeks after the respective election day.[5] In the surveys, the specific question for the winner was worded as follows: "In your opinion, which of the four politicians fared best overall in this debate?"

The parliamentary nature of the West German political system, the coalitional structure of the government, and the multiparty nature of the system prevent a head-to-head meeting of the two chancellor candidates. Instead, chancellor candidates and leaders of the parties repre-

---

[2] In the 1980 election survey, sociodemographic control variables and partisanship were included due to the lack of panel data.

[3] The National German Election Surveys were made available by ICPSR (study numbers 7102, 7513, 7963), and the Zentralarchiv für Empirische Sozialforschung, Universität Köln (ZA-Nr. 1275 and 1276). All surveys are probability samples of the national population. These studies were also used for the subsequent analyses in this research project.

[4] I intended to include the analysis of the most recent debate of 1987. Unfortunately, no data was gathered for that debate and, hence, the analysis could not be extended beyond the election of 1983.

[5] Since debates do not seem to attract the interest of West German journalists, practically no postdebate analyses—as in the United States—exist. Debates, their content, or their outcome are not discussed in newspapers or on television.

Table 1. Participating Politicians in Debates, 1972–1983

Party[a]	1972	1976	1980	1983
SPD	Brandt	Schmidt	Schmidt	Vogel
CDU	Barzel	Kohl	Kohl	Kohl
FDP	Scheel	Genscher	Genscher	Genscher
CSU	Strauss	Strauss	Strauss	Strauss

[a] Party abbreviations refer to the Social Democratic Party (SPD—Sozialdemokratische Partei Deutschland), the Christian Democratic Party (CDU—Christlich Demokratische Union Deutschlands), the Christian Social Party (CSU—Christlich Soziale Union Deutschlands), and the Free Democratic Party (F.D.P.—Freie Demokratische Partei).

sented in the Bundestag participate in West German debates.[6] The participants in the 1972, 1976, 1980, and 1983 debates were the party leaders of the FDP, the CDU, and the CSU, and the chancellor candidate of the SPD (the incumbent chancellor in 1972, 1976, and 1980). The chancellor candidate of the political opponents was typically the party leader of the CDU, except in 1980, when CSU leader Strauss was the candidate of the CDU/CSU (table 1).

The goal of this analysis was to test possible attitudinal, affective, and behavioral effects. Since it is conceivable that the audience perceives a difference between the politicians (Lang and Lang, 1979; Miller and MacKuen, 1979), the study will show to what degree the debate in general, and the perception of the winner in particular, affects the individual spectator.

Three models were tested: the first probes for a general debate impact on the individual politician's evaluation; the second tests the effect of the perception of the winner on the spectator's evaluation; and the final test establishes the effect of the perception of the winner on the vote decision. In the first and all subsequent models, each politician's evaluation (PE) is measured in the pre- and postdebate questionnaire with an eleven-point scalometer question, where the respondent was asked to rate the politician from "Do not think much of the politician at all" (coded 1) to "Think a great deal of the politician" (coded 11). The debate viewer variable is a dummy variable (DEB) that distinguishes between not watching the debate and watching parts of or the entire debate. The winner variable is a dummy variable scored 1 if a participant is depicted as the winner and 0 if all participants were rated "equally well/bad." For the final model, the dependent variable is the vote cast (V) for the Government in power. It consists of a dummy variable, scored 1 if the respondent cast his or her vote for one of the government parties and 0 if the vote is cast for an opposition party.

---

[6] Until 1983 there were four parties in the Bundestag: CDU, CSU, FDP, and SPD. Since the 1983 election a fifth party, the Greens, overcame the 5% vote share barrier to obtain seats in the Bundestag.

Party identification (PID) and the politicians' evaluations (PE) are included as control variables because they have been shown to determine the vote to a large extent (Campbell et al., 1960; for Germany see Norpoth, 1977; Falter and Rattinger, 1983). Party identification ranges from 1 (strong SPD-FDP in 1972, 1976, 1980, and only SPD in 1983) to 5 (strong CDU/CSU in 1972, 1976, 1980, and CDU/CSU-FDP in 1983). This represents the partisanship for the parties being either the Government coalition or the opposition parties.

The parameters of the evaluation and winner models are estimated through the use of ordinary least squares. To avoid spurious relationships and ascertain debate effects in these models, the predebate evaluations ($PEt - 1$) are included to control for the predispositions of individuals. The inclusion of this control variable allows us to separate debate effects from those of other factors such as partisanship and demographic variables.[7] Since it is inappropriate to use ordinary least squares when dealing with a dichotomous dependent variable (Aldrich and Nelson, 1984), the coefficients of the vote models were estimated by using probit analysis, a maximum likelihood estimation technique.

Results

EXPOSURE TO THE DEBATE

In each election year a majority of the electorate watched at least parts of the debates. The three debates of 1972 were viewed by an overwhelming 84% of the electorate. The 1976 debate, this time a single event, drew the interest of 75%. In 1980 it was 68%, and in 1983, 56% of survey respondents said they watched the debates. Although this indicates a decline in interest in the televised debates, it still signifies that during each election campaign over half of the electorate was exposed to parts of or all of the debates. These audience numbers suggest that more people watched the debates than are politically very interested or claim to be strong partisans. It thus tells us that effects can possibly be expected, and underlines the importance of such an event for West German elections, especially since the debates take

---

[7] In 1980 age (AGE), occupation (OCC), and education (EDUC) are included because, unlike the other data sets, the election survey consisted of cross-sectional data only.

Table 2. Evaluations of Debate Participants (PE) by Watching Debates (DEB) (Standardized regression coefficients; standard errors in parentheses)

Debate	SPD Party Leader	CDU Party Leader	FDP Party Leader	CSU Party Leader
1972	.007	−.016	NA	NA
	(.15)	(.20)		
1976	−.042*	−.045**	.038	−.035*
	(.13)	(.13)	(.12)	(.15)
1980	.078***	−.038	.077***	−.092***
	(.28)	(.29)	(.33)	(.24)
1983	−.053**	.055**	.006	.036
	(.12)	(.12)	(.14)	(.14)

Note: Predebate evaluations were held constant in the estimation of the debate effect. NA = not applicable.
$^* p < .05$. $^{**} p < .01$. $^{***} p < .001$.

place only a few days prior to Election Sunday. Therefore, whatever the viewer learns, and whatever his or her impressions are, it will be fresh in short-term memory when the time comes to cast the vote.

WATCHING THE DEBATES AND EVALUATIONS OF POLITICIANS

The first test attempted to answer the question of whether watching the debates (DEB) influenced postelection evaluation of the participating politicians (PE). Table 2 shows that, except for 1972, debate watching makes a significant difference for one or more of the participating candidates. A positive sign of the debate variable (DEB) means that debate watchers rated that candidate higher than nonwatchers, while a negative sign implies a lower rating for viewers. In 1972, there is no difference between debate watchers and nonwatchers in terms of their evaluations of the chancellor or his challenger. However, in that year, instead of a single debate prior to the election, a series of three debates over a time period of several weeks was conducted. Moreover, the question asked was whether the respondent watched one, two, three, or none of the debates. About 84% of the respondents said they watched at least one of the debates, which makes it difficult to distinguish attitudinal changes of debate watchers from those of nonwatchers. Furthermore, the debate viewer variable contains respondents who watched perhaps only a few minutes and those who watched several hours, and it is technically not possible to separate them.

The three debates were also different in length and format. As Baker and Norpoth (1981) report, the first two debates, which were shorter than the third, were more closely supervised by journalists who asked questions. The final debate, in contrast, was similar to the debates in the following years, which means that it was held very shortly before the election and that it was characterized by a strong interaction among the party leaders while the journalists played only a minor role. Thus a positive effect after the first debate due to a positive display of competence might have been offset by a weak showing in direct confrontation. Unfortunately, without the possibility of testing every debate individually there is no way of finding out whether or not watching the third debate changed candidate evaluations.[8]

In 1976, debate watching made a significant difference for three out of the four politicians. In all three cases, however, the effect was negative. Chancellor Schmidt (SPD), his challenger Kohl (CDU), and CSU leader Strauss were rated more negatively afterwards by debate watchers than by those who did not watch. For Genscher (FDP), the estimate is positive but insignificant. Thus it is not possible to say that he gained from the debate.

The findings for 1980 again show a debate effect for three out of the four participants: Chancellor Schmidt and his coalition partner Genscher (FDP) both registered positive gains, and the chancellor challenger Strauss (CSU) showed an equally strong negative impact. The estimate for Kohl was insignificant.

The results for the debate in 1983 show a positive debate effect for chancellor Kohl (CDU), and a negative effect for his challenger Vogel from the SPD. The debate did not add to or subtract significantly from the evaluation of the government coalition partners (Genscher from the FDP and Strauss of the CSU).

Table 2 contains two noteworthy findings. First, apart from 1972, watching the debate always contributed significantly to the ratings of the chancellor and his challenger, while only rarely changing the evaluation of the chancellor candidates' coalition partners. Second, the debate effect is not consistently positive or negative with respect to the incumbent chancellor, thus failing to support a possible incumbency effect—*Kanzlerbonus* (Noelle-Neumann, 1983). The chancellor challenger, however, *always* loses ground among debate viewers. While the chancellor managed to receive a boost in 1980 and 1983, his challenger always fared worse and lost through the debate. This "challenger effect" might be a reflection of the strategy applied. That is to say, challengers usually tend to pursue a negative debate style by attacking their opponent (Schrott, 1989). The intention for such a strategy is to overcome a possible incumbency effect. By attacking an incumbent office holder, differences between perceptions of the electorate and the actual performance in office can be shown and used against the incumbent (Martel, 1983).

---

[8] Tests were performed to see whether the number of debates watched made a difference, but results suggest it did not matter whether the respondents watched one, two, or all three of the debates. Furthermore, the same model was tested using the politician evaluation questions of the first and second wave that were done right before and after the first debate. The debate variable still proved to be insignificant. But again, the problem was that the debate variable included *all* watchers, no matter whether they watched only the first or the second and/or the third. Nevertheless, the test was performed because Baker and Norpoth (1981) found some differences concerning certain issues. Yet, in terms of politician evaluations no difference between watchers and nonwatchers could be established.

Interestingly, these results deviate from the U.S. setting, where debates seem to help mainly the challenger and less the incumbent president (Kraus, 1962 and 1979). In the West German setting the reverse seems to be true; it is the challenger who constantly loses ground, and the chancellor who appears to gain. Again, the explanation may be found in different debate strategies. The challenger in the United States is often a relatively unknown politician—such as John F. Kennedy in 1960—who may use the opportunity to present himself as a competent politician (Kraus, 1962). The West German debaters, on the other hand, are widely known politicians (Norpoth, 1977). Thus, the debates seem to reflect the parliamentary confrontation between the opposition and government insofar that the challenger attacks while the chancellor defends, possibly leaving a negative image for the challenger (Norpoth and Baker, 1983).

THE IMPACT OF WINNING DEBATES ON CANDIDATE EVALUATIONS

The next step is to incorporate the individual's perception of the winner in the model of politician evaluation to see whether the winner becomes more positively evaluated as a result of his performance. A viewer's choice of the debate winner was included as a dummy variable in the regressions for each of the four participating party leaders. If a respondent picked the SPD party leader as the winner, the variable "SPD Winner" was coded 1, 0 if not. The same procedure was applied for each of the other participants. The question about debate victory also allowed respondents to answer "All four fared equally well" if they could not decide to choose a winner. This option was also included as the dummy variable "All equal." The behavior of the nonwatchers is captured in the constant term. An insignificant winner estimate for a particular politician would indicate that his rating did not change as the result of his being chosen as the winner. The results are displayed in table 3.[9]

Table 3 shows that a candidate who is picked as the winner of the debate always receives a positive boost from that part of the audience. This holds for all politicians in the debate. Table 3 also indicates that effects are seen for the *other* participants when a given politician is perceived as the winner, that is, the other candidates register declines in their overall evaluation. The magnitude of the estimates is rather large, translating into a gain in evaluation from one-half to one and one-half points on an eleven-point scale. Clearly, to be rated as the winner helps a politician enormously in his evaluation.

In 1972, for example, for Chancellor Brandt to be judged the winner translated roughly into a one unit increase in his evaluation. However, if a debate viewer selected the CDU politician as the winner, he rated Brandt about one unit lower. Thus, there was a difference of roughly two units between the respondents who rated the chancellor or his challenger as the winner. Favoring the CSU politician as the winner meant an even sharper drop in Brandt's rating. Yet, a respondent favoring the FDP leader still rated Brandt more than a half-unit higher. These findings indicate that "coalitional" relations translate into favorable politician evaluations. As long as the SPD and the FDP were coalition partners, respondents evaluated the party leaders positively if either of them were judged as the winner.

The inverse holds for the opposition. Every time an opposition leader was chosen as the winner, the impact on the government party leaders was negative, while the impact on the leader of the other opposition party was positive. Still, the impact through the leader of the smaller opposition party is, in some cases, statistically insignificant. In 1972, the choice of Strauss (CSU) as the winner did not increase Barzel's (CDU) evaluation. The same happened in 1976 to Kohl (CDU). People who rated Strauss as the winner failed to evaluate their party's (CDU/CSU) chancellor candidate more positively. The same people, however, clearly disliked the SPD and FDP politicians. In 1983, the respondents who thought that Strauss fared best also rated Chancellor Kohl more favorably. Interestingly enough, the people who saw Genscher, FDP, as the winner did not change their evaluation, and thus did not display any "coalitional" bias.[10] (The estimates for the other politicians are nevertheless in the "right" direction, i.e., negative for the SPD politician and positive for the CDU/CSU party leaders.)

---

[9] Again, no estimation was possible in 1980 because no winner question was asked in that survey.

[10] By 1983 the FDP was in a government coalition with the CDU/CSU.

Table 3. Evaluations of Debate Participants by Winning Debates (Unstandardized regression coefficients; standard errors in parentheses)

	1972		1976				1983			
	Brandt (SPD)	Barzel (CDU)	Schmidt (SPD)	Kohl (CDU)	Genscher (FDP)	Strauss (CSU)	Vogel (SPD)	Kohl (CDU)	Genscher (FDP)	Strauss (CSU)
PEt − 1	.600***	.530***	.586***	.586***	.453***	.619***	.566***	.599***	.566***	.609***
	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)	(.03)	(.02)
SPD winner	.972***	−1.128***	.847***	−1.250***	.509***	−1.492***	.579***	−.766***	−1.080***	−.736
	(.15)	(.20)	(.17)	(.17)	(.16)	(.21)	(.18)	(.18)	(.22)	(.22)
CDU winner	−.878***	1.354***	−1.545***	1.045***	−.656***	.950***	−.974***	1.140***	.685***	.768***
	(.20)	(.26)	(.17)	(.17)	(.16)	(.21)	(.15)	(.15)	(.18)	(.18)
FDP winner	.592***	−.753***	.101	−.607***	.844***	−.749***	−.366	.390	1.265***	.122
	(.16)	(.21)	(.15)	(.15)	(.14)	(.18)	(.28)	(.28)	(.34)	(.34)
CSU winner	−1.176***	.696	−1.389***	.287	−.422**	1.019***	−.659***	.487**	.044	1.092***
	(.24)	(.31)	(.26)	(.25)	(.25)	(.31)	(.29)	(.29)	(.35)	(.36)
All equal	−.060	.497*	−.300*	.030	−.401***	.170	.028	.088	−.280	.221
	(.23)	(.30)	(.19)	(.19)	(.18)	(.23)	(.19)	(.19)	(.23)	(.23)
Constant	3.364	2.979	3.463	3.448	4.111	2.723	3.180	3.357	3.167	2.452
R^2	.60	.48	.56	.54	.36	.58	.50	.56	.41	.52
N	1,212	1,211	1,177	1,172	1,167	1,174	1,002	1,006	1,004	1,005

* $p < .05$. ** $p < .01$. *** $p < .001$.

Thus, it appears that being the winner helps a politician substantially to increase his rating. At the same time, it hurts the rating of opposing politicians. This is important in light of research that found personal evaluations to be one of the determining factors of the vote choice (Klingemann and Taylor, 1977; Norpoth, 1977; for the United States, see Campbell et al., 1960).

DEBATES AND THE VOTE DECISION

The final step of the analysis concentrates on testing whether the debates had any behavioral effects upon the electorate. Did watching the debates influence the members of the audience in their vote decision? If the debates provided new insight into political matters so that the audience altered its vote intention, or if the debate viewers were so impressed by a certain politician that they voted for his party contrary to their own party affiliations, the conclusion could be drawn that the debates were an important political event in influencing the election. The winner judgment could influence the vote choice in several ways: first, through the impact on candidate evaluations that was shown above, and second, as an independent cue for voting based on the debates' extremely close proximity to the elections.[11] The individual vote decision is treated as a dummy variable, with a vote in favor of the incumbent government parties scored 1, and a vote for other parties scored 0. The parameters are estimated using probit analysis, a maximum-likelihood estimation technique. The MLE's represent the estimate of the change in the cumulative normal probability function given a one unit change in the independent variable. The test statistic for the significance of the parameter estimates is MLE/SE, and the goodness of fit is measured by the likelihood ratio principle (Aldrich and Cnudde, 1975).

The analysis proceeds in two steps. First, a vote model incorporating party identification (PID) and the evaluation of the politicians is estimated; then, variables for the effects of the judgment of the debate winner are added, and the models are compared. The results for the basic vote model are reported in table 4.

The estimates for party identification are highly significant and very large in magnitude. A similar picture is drawn with respect to the estimates for the evaluation of the "candidates," meaning the chancellor, the chancellor candidate of the opposition, and the leaders of the

---

[11] Issue variables were included in the original model but then dropped because of statistical insignificance. That is not to say that issues were of no importance in the elections. It only means that the selected issues of economic performance were statistically insignificant (for the reasoning that had led to the use of these variables, see Roth, 1973).

Table 4. Vote Choice: Probit Estimates (MLE/SE in parentheses)

	1972	1976	1980	1983
Party identification	−.756***	−.298***	−.434***	.174***
	(−12.90)	(−6.15)	(−5.89)	(4.57)
Predebate evaluation				
SPD leader	.191***	.250***	.251***	−.267***
	(6.74)	(7.93)	(5.97)	(−10.47)
CDU leader	−.103***	−.252***	−.321***	.143***
	(−4.10)	(−7.70)	(−7.51)	(4.73)
FDP leader	NA	.103**	.232***	−.011
		(3.02)	(4.96)	(−.45)
CSU leader	NA	−.146***	−.230***	.106***
		(−6.72)	(−7.55)	(4.70)
Constant	1.591	1.161	1.888	−.090
$\bar{y}$	.64	.54	.65	.53
−2.0 log	945.80	843.59	716.53	548.08
df	3	5	5	5
Percentage predicted correctly	.94	.90	.92	.85

NOTE: NA = not applicable.
$* p < .05$. $** p < .01$. $*** p < .001$.

allied partner (FDP, CSU). The MLE/SE values for the estimate of the SPD leader (the chancellor, 1972–80) increases from 1972 to 1976, is lowest in 1980, and highest in 1983. For the other politicians as well these values vary from election to election. This does not support the thesis that a decrease in the importance of partisanship is accompanied by an increase in the importance of the politician evaluation. It rather appears that politician evaluation fluctuates in its importance for the vote. The estimate for Strauss (CSU), is the largest in 1980, when he was the chancellor candidate of the opposition. The estimates for Kohl (CDU) were strongest when he was chancellor candidate, in 1976, and when he was chancellor, in 1983. That seems to indicate that personal evaluations matter most for chancellor candidates, a notion that is also supported by the estimate for SPD chancellor Brandt in 1976, and Vogel, the chancellor candidate of the opposition in 1983. As their evaluations decrease, voters are more likely to vote for the opposing party or parties.

Adding the winner variables increased the explanatory power of the vote model quite substantially in two of the three instances, as indicated by the significantly larger chi-square shown in table 5. Even in 1972, the improvement of the chi-square is too close to the .05 level

Table 5. Winning the Debates and Vote Choice: Probit Estimates (MLE/SE in parentheses)

	1972	1976	1983
Party identification	−.725***	−.346***	.151***
	(−12.12)	(−5.94)	(3.81)
SPD winner	.120	1.335***	−.420*
	(.64)	(5.10)	(−2.14)
CDU winner	−.443	−1.43***	.819***
	(−1.79)	(5.77)	(5.41)
FDP winner	.046	.463**	.573*
	(.24)	(2.85)	(2.07)
CSU winner	−.695*	−1.097*	.803*
	(−1.96)	(−2.55)	(2.51)
None	−.372	−.540*	.322
	(−1.39)	(−2.45)	(1.86)
Predebate evaluation			
SPD leader	.171***	.244***	−.253***
	(5.80)	(6.85)	(−9.47)
CDU leader	−.090***	−.251***	.131***
	(−3.49)	(−6.83)	(4.15)
FDP leader	NA	.104**	−.013
		(2.65)	(−.48)
CSU leader	NA	−.144***	.106***
		(−5.81)	(4.49)
Constant	1.641	1.345	−.239
$\bar{y}$	.64	.54	.53
−2.0 log	956.27	984.75	601.18
df	8	10	10
Percentage predicted correctly	.94	.92	.85

NOTE: NA = not applicable.
* $p < .05$. ** $p < .01$. *** $p < .001$.

to discard the winner perception as a possible cause of the vote decision. In 1976 and 1983, the chi-square improvement is significant on the .001 level. The debate of 1976 shows large and significant parameter estimates for all winner variables, including the option that no winner was chosen. For Chancellor Schmidt, SPD, to be seen as the winner increased the probability that the vote was cast for the Government quite substantially (MLE of 1.335). The government parties gained also if his coalition partner from the FDP was judged the winner, as indicated by the estimate of .463. The largest single winner estimate was − 1.43 for the CDU participant, and thus, decreased the probability to vote for the government parties to a large extent. People who saw Strauss as the winner were also significantly less likely to vote for the government (MLE of 1.097). And last, if people saw all equal, their likelihood to vote for the government decreased. The estimate of − .54 is rather large and significant, and might indicate that people were generally dissatisfied with the government leaders and turned against them. Clearly, losing the debate hurt the incumbent politicians and not the challengers.

The latest debate in 1983 displayed similar but somewhat weaker results than the 1976 contest. Being selected the winner helped the party of each participant. The largest estimate in the 1983 model is for the CDU winner (.819); in this case the probability to vote for one of the CDU, CSU, FDP government parties increases. A similarly strong estimate is found for the CSU winner (.803), and even for the FDP leader (.573). People failing to perceive a winner were still more likely to vote for the government. This time it did not hurt the government parties if their leaders were not seen as winner.

The weak estimates in 1972, especially for Chancellor Brandt and his challenger, Barzel, might be explained by the nature of the 1972 debates. As pointed out earlier, the 1972 debates consisted of a whole series presented over several weeks. This might have led the electorate to change its attitudes from one debate to the next, and somebody they liked in the first debate they might have disliked in the second debate, and so on. Or their first impression was changed through the media offering a "winner"; or by the time they were asked to pick a winner they simply opted for the "mediated" winner (Lang and Lang, 1979).

The effects for the various winner variables tell us several interesting facts. The magnitude of these estimates vary from election to election and from winner to winner, but the overall picture is one of importance in determining the vote. In nine out of twelve cases, the estimates for the winner are quite substantial, and are significantly related to the probability of voting for the governing parties. Not only the chancellor and his challenger, but the other participants, too, are likely to gain votes for their parties if judged as the winner. This indicates that the debates in West Germany are not simply a contest between the chancellor and his challenger but involve the other two participants equally strongly. A small party such as the FDP might gain substantially from its leader being the winner of the debate.

The overall fit of the model suggests that the inclusion of the winner variables adds substantially. This is shown in the improvement of the chi-squares for the nested models. These findings underline the crucial importance for the participants to perform well in the debates and to be chosen as the winner. For the chancellor and his challenger to be judged the winner helps them sway undecided voters to their partisan side. The same holds for the small FDP and its party leader. At least in 1976, Genscher's debate performance was successful and he increased the vote share of the SPD-FDP parties. Of course, these findings do not tell us whether these people cast their vote for the FDP or the SPD. A benefit for the FDP seems likely, however, because Genscher's debate performance was aimed at selling himself and his party rather than the governing coalition as such (Schrott, 1986). For a small party that has to compete with all the other parties, it is especially crucial that its party leader performs well, so that it can be sure to reach the minimum 5% barrier that is necessary to obtain seats in the Bundestag. A party like the CSU does not have the same problem, since it is the only representative of the CDU/CSU coalition in Bavaria. That does not imply, of course, that for its party leader to be judged the winner would not increase the CDU/CSU's vote share.

Summary and Conclusion

This research has examined how televised debates in West Germany affect the audience. The West German case provided an ideal opportunity to study the effects on the audience over a substantial period of time. Since 1972, televised debates represent a common ingredient in the Federal Republic's electoral campaign. Each national election in 1972, 1976, 1980, 1983, and 1987 was preceded by at least one televised debate. The results show unequivocally that the debates affected public evaluations of the participants and vote choice in almost all elections since 1972. Debate viewers changed their evaluations of the chancellor and his challenger in all but one election year. In the Federal Republic, it appears that both contenders for the major political office are watched and evaluated rather carefully. The challenger always suffered in the debate, thus pointing to a negative "challenger effect." The incumbent chancellor, on the other hand, gained in two out of three debates; the estimate for 1972, again, while positive, is insignificant. It appears that the chancellor is more likely to gain from a debate than his challenger. With respect to the audience's evaluation of the participants' performance, the winner always received a boost in his evaluation, and in most models the perceived winner had a negative impact on the other politicians.

Finally, the inclusion of a variable measuring voter judgment of the debate winner in a basic model of vote choice increased each model's explanatory power significantly. This effect was especially large if the chancellor was seen as the winner. The likelihood of a government party receiving the respondent's vote went up considerably in that event. The same held for the opposition's chancellor challenger. If he was the perceived winner, the likelihood of voting for a government party decreased.

In sum, this study provides strong evidence suggesting that televised debates in West Germany provide the audience with some important cues on which to base the vote. Not only is there a direct impact for the perceived winner of the debates, but there is also an indirect effect, in that a politician gains through the debates in his personal evaluation, which in turn helps his party's electoral support. The debate may be of greatest importance for undecided and independent voters who might vote for the perceived winner due to their lack of a strong identification with a party or a politician. The question remains, of course, as to how likely it is that people are undecided at the time of the debate or did not have a preferred party or chancellor candidate. However, in close elections, the small percentage of people voting for the perceived winner might decide the political future of the country. And considering the fact that the debates in West Germany are held a few days prior to election day, they gain even greater importance because the possible short-term effects (Miller and MacKuen, 1979) might last just through election day, and the winner of the debate might be the most salient memory for some of the voters.

References

Abramowitz, Alan I. (1978)
"The impact of a presidential debate on voter rationality." American Journal of Political Science 22:680–90.

Aldrich, John, and Forrest Nelson (1984)
Linear Probability, Logit and Probit Models. Beverly Hills: Sage Publications.

Baker, Kendall L., and Helmut Norpoth (1981)
"Candidates on television: The 1972 electoral debates in West Germany." Public Opinion Quarterly 45:329–45.

Baker, Kendall L., Helmut Norpoth, and Klaus Schönbach (1981)
"Die Fernsehdebatten der Spitzenkandidaten vor den Bundestagswahlen 1972 und 1976." Publizistik 26:530–44.

Becker, Lee B., Idowu A. Sobowale, Robin E. Cobbey, and Chaim H. Eyal (1980)
"Debates' effects on voters' understanding of candidates and issues." In George F. Bishop, Robert G. Meadow, and Marilyn Jackson-Beeck (eds.), The Presidential Debates. New York: Praeger.

Bishop, George F., Robert G. Meadow, and Marilyn Jackson-Beeck, eds. (1978)
The Presidential Debates. New York: Praeger.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes (1960)
The American Voter. New York: Wiley.

Falter, Jürgen, and H. Rattinger (1983)
"Parteien, Kandidaten und politische Streitfragen bei der Bundestagswahl 1980: Möglichkeiten und Grenzen der Normal-Vote-Analyse." In Max Kaase and Hans-

Dieter Klingemann (eds.), Wahlen und politisches System. Opladen: Westdeutscher Verlag.

Graber, Doris A. (1978)
"Problems in measuring audience effects of the 1976 debates." In George F. Bishop, Robert G. Meadow, and Marilyn Jackson-Beeck (eds.), The Presidential Debates. New York: Praeger.

Kaase, Max, and Hans-Dieter Klingemann, eds. (1983)
Wahlen und politisches System. Opladen: Westdeutscher Verlag.

Katz, E., and J. J. Feldman (1962)
"The debates in the light of research: A survey of surveys." In Sidney Kraus (ed.), The Great Debates. Bloomington: Indiana University Press.

Klingemann, Hans-Dieter, and C. L. Taylor (1977)
"Affektive Parteiorientierung, Kanzlerkandidaten und Issues." Politische Vierteljahresschrift 18:301–47.

Kraus, Sidney, ed. (1962)
The Great Debates. Bloomington: Indiana University Press.

——— (1979)
The Great Debates: Carter vs. Ford, 1976. Bloomington: Indiana University Press.

Kraus, S., and D. Davis (1976)
The Effects of Mass Communication on Political Behavior. University Park, PA: Pennsylvania State University Press.

Lang, Kurt, and Gladys Engel Lang (1962)
"Ordeal by debate: Viewer reactions." Public Opinion Quarterly 25:277–88.

——— (1979)
"Immediate and mediated responses: First debate." In Sidney Kraus (ed.), The Great Debates: Carter vs. Ford, 1976. Bloomington: Indiana University Press.

McLeod, J. M., J. A. Durall, D. A. Ziemke, and C. R. Bybee (1979)
"Reactions of young and older voters: Expanding the context of effects." In Sidney Kraus (ed.), The Great Debates: Carter vs. Ford, 1976. Bloomington: Indiana University Press.

Martel, Myles (1983)
Political Campaign Debates Images, Strategies, and Tactics. New York: Longman.

Miller, A. H., and Michael MacKuen (1979)
"Informing the electorate: A national study." In Sidney Kraus (ed.), The Great Debates: Carter vs. Ford, 1976. Bloomington: Indiana University Press.

Noelle-Neumann, Elisabeth (1983)
"Öffentliche Meinung in der Bundestagswahl 1980." In Max Kaase and Hans-Dieter Klingemann (eds.), Wahlen und politisches System. Opladen: Westdeutscher Verlag.

Norpoth, H. (1977)
"Kanzlerkandidaten." Politische Vierteljahresschrift 18:551–72.

Norpoth, Helmut, and Kendall L. Baker (1983)
"Politiker unter sich am Bildschirm: Die Konfrontation von Personen und Sachthemen in den Fernsehdiskussionen 1972–1980." In Max Kaase and Hans-Dieter Klingemann (eds.), Wahlen und politisches System. Opladen: Westdeutscher Verlag.

O'Keefe, Garrett J., and Harold Mendelsohn (1979)
"Media influences and their anticipation." In Sidney Kraus (ed.), The Great Debates: Carter vs. Ford, 1976. Bloomington: Indiana University Press.

Ranney, Austin, ed. (1979)
The Past and the Future of Presidential Debates. Washington, DC: American Enterprise Institute.

Roth, Dieter (1973)
"Ökonomische Variablen und Wahlverhalten." Politische Vierteljahresschrift 14:257–72.

Schrott, Peter R. (1986)
"The West German television debates, 1972–1983: Candidate strategies and voter response." Ph.D dissertation. State University of New York, Stony Brook.

——— (1989)
"The West German television debates, 1972–1987: Candidate strategies and voter

response." Paper delivered at the annual meeting of the American Political Science Association, Atlanta, GA.

Sears, D. O., and S. H. Chaffee (1979)
"Uses and effects of the 1976 debates: An overview of empirical studies." In Sidney Kraus (ed.), The Great Debates: Carter vs. Ford, 1976. Bloomington: Indiana University Press.

Weiss, Hans-Jürgen (1976)
Wahlkampf im Fernsehen: Untersuchung zur Rolle der grossen Fernsehdebatten im Bundestagswahlkampf 1972. Berlin: Verlag Volker Spiess.

White, D. M. (1950)
"The gate-keeper: A case study in the selection of news." Journalism Quarterly 27:383–90.

THE IMPACT OF TELEVISION VIEWING ON MENTAL APTITUDE AND ACHIEVEMENT: A LONGITUDINAL STUDY

STEVEN L. GORTMAKER,
CHARLES A. SALTER,
DEBORAH K. WALKER, and
WILLIAM H. DIETZ, JR.

Abstract Previous studies of the effects of television viewing on mental abilities have shown mixed results, but most suffered from one or more of the following shortcomings: a small or otherwise unrepresentative sample, a cross-sectional rather than longitudinal approach, and a failure to consider intervening variables between television viewing and cognitive skills. This study was designed to overcome these deficiencies by using nationally representative data from the National Health Examination Cycle 2 and Cycle 3 Surveys. These surveys included 1,745 children who were studied both when 6 to 11 years old (between July, 1963, and December, 1965) and about four years later when 12 to 17 years old (between March, 1966, and March, 1970). Simple analysis at just one time-point reveals substantial relationships between the amount of television viewing and depressed IQ and Wide Range Achievement Test (Reading and Arithmetic) scores of adolescents. When longitudinal controls are added, however, these relationships become statistically insignificant and substantively unimportant. Although these data are 20 years old, they indicate no significant causal relationship between the amount of television viewed and the mental aptitude and achievement test scores of adolescents, thus supporting and extending Gaddy's (1986) recent longitudinal study of a national sample of youth.

STEVEN L. GORTMAKER is Senior Lecturer and Acting Chairman of the Department of Behavioral Sciences at Harvard University School of Public Health. CHARLES A. SALTER is a Health Services Research Psychologist at Natick (MA) Research, Development & Engineering Center. DEBORAH K. WALKER is a Lecturer in the Department of Behavioral Sciences at Harvard University School of Public Health. WILLIAM H. DIETZ, JR., is Chief of the Department of Pediatric Gastroenterology and Nutrition at Tufts New England Medical Center. The research presented here was supported by Grant #85103185 from the William T. Grant Foundation.

Public Opinion Quarterly 54:594–604 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-04/$2.50

Never in history has there been a mass medium more omnipresent than television. The average family member in the United States watches television for about half of his or her nonsleeping hours at home (McLeod et al., 1982). In November, 1984, the average preschooler watched 28 hours and 20 minutes per week, while the average 6- to 11-year-old watched 26 hours and 34 minutes per week (World Almanac, 1986). These figures indicate that the average child spends more hours per year watching television than attending school. Just as schooling is presumed to have profound cognitive effects, one might assume that such a heavy, long-term diet of television could have major effects on children's cognitive abilities.

But what kind of effects should be expected? As Hornik (1978) pointed out, there are several possibilities. The consequences of television viewing could be negative, with the television displacing time from more beneficial activities such as reading, hobbies, artistic and creative skills, or imaginative play (Neuman, 1988). On the other hand, the consequences of viewing could be positive. Television not only may entertain, but may inform and stimulate as well. It provides ready access to worldwide news, cultural events, and positive role models. It may stimulate interest in academically relevant topics to which a child may not otherwise be exposed. Other writers (Rubinstein, 1983) have theorized that television may have differential effects, taking time away from those who live in beneficial environments, thus inhibiting their mental development, but adding new dimensions of intellectual stimulation to those children in deprived environments, thus improving their cognitive skills. Several previous studies have suggested a negative effect of television (e.g., Himmelweit, Oppenheim, and Vince, 1958; Thompson, 1964; Witty, 1967; Singer, 1980; Williams et al., 1982). Other studies have found a positive effect (e.g., Greenstein, 1954; Schramm, Lyle, and Parker, 1961), while Hornik (1978) found no effect over the short term, but a negative effect on reading improvement over time. However, these previous studies generally had at least one of the following limitations: a small or otherwise unrepresentative sample, study at only one time-point, or a failure to control for other variables in addition to television viewing and mental achievement.

A recent study by Gaddy (1986) did examine a large, randomly selected, national sample of U.S. teenagers, examining them at two points in time: as high school sophomores and two years later as seniors. This study showed that simple correlations between television viewing and achievement suggest a negative effect, but when other variables are included in the analysis, this negative association becomes insignificant.

The purpose of the present study is to confirm and extend Gaddy's findings by examining a large, random, national sample of U.S. youth at two different times, but starting at a younger age and following them for about four to six, rather than two, years. It is possible that effects of television viewing would show up with younger children over a longer period of observation.

Method

SUBJECTS

The subjects in this study were 1,745 children who were included in both Cycle 2 and Cycle 3 of the National Health Examination Surveys (NHES), which occurred in 1963–65 and 1966–70, and for whom complete data on all variables was obtained. The children were aged 6–11 in Cycle 2 and 12–17 in Cycle 3. The sampling plan resulted in a representative sample of the entire noninstitutionalized population of the United States for each age group studied (National Center for Health Statistics, 1973). The response rate for the Cycle 2 survey was 96%, and the response rate for the Cycle 3 survey was 90%.

MATERIALS AND PROCEDURES

The National Health Examination Survey included in Cycle 2 a parental report of the amount of television viewed daily, while Cycle 3 included a self-report on this behavior. The question in Cycle 2 was, "About how much time would you guess your child spends on the usual day doing each of the following: (Enter number of hours or fraction of hours or zero as appropriate) a. Watching television? _____." The question in Cycle 3 was, "About how much time would you guess you spend in the usual day (*enter number of hours or fraction of hours, or zero, as appropriate*)? a. Watching television _____." In both cases, reports were coded into half-hour intervals for up to two hours, and then hour-long intervals after that. The mean of hours of television viewed daily was calculated from the midpoint of each hourly time-interval except for the last interval, when an additional half-hour was added to account for the outer tail of the distribution. The reported behavior is thus retrospective—the parent or child is asked to refer to the recent past—while the tests were administered at roughly the same time as the behavioral questionnaires. Anderson et al. (1985) found that such parental reports of children's television viewing did correlate (r = .48) with parental estimates of viewing obtained via ten-day daily diaries. These diary estimates have, in turn, been shown to be strongly related to actual viewing as measured by time-lapse video home observation. And Dietz and Gortmaker (1985) found that the same television viewing measures as used in this study did significantly predict childhood obesity independent of a wide variety of socioeconomic variables. Furthermore, Ritchie, Price, and Roberts (1987) found a high stability in self-reported television watching among children over time, with stability coefficients over three years ranging from .72 ($p < .001$) to .85.

Cognitive skills were assessed at both times by the Wechsler Intelligence Scale for Children (WISC), by the Wide Range Achievement Test in Arithmetic (WRAT-A), and the WRAT in Reading (WRAT-R). Normalized standard scores of the WISC vocabulary and block design *t*-scores were summed to create an Index of Intellectual Development. Normalized standard WISC vocabulary, WISC block design, WRAT-R, and WRAT-A scores were based upon the estimated population distribution of children examined in the Cycle 2 and Cycle 3 surveys. Scores were normalized by four-month age-intervals and adjusted for sex differences. Testing was carried out by Master's level psychologists with previous experience in test administration (National Center for Health Statistics, 1966).

In addition to these two types of variables of primary interest—television viewing and cognitive abilities—the NHES included a large number of variables that needed to be controlled in examining the relationship between television viewing and cognitive abilities and achievement. Included in this analysis were size of place of residence, region of the country, season of data collection, birth order, race, family income, parental education, the presence of a physically disabling condition, and parental age. It was important to control for both season and region because the examinations were conducted in the North during the summer and in the South during the winter. Dummy variables were used to represent region, season, and size of place, with the omitted category being indicated in table 4. For a complete discussion of which variables were assessed and how, see National Center for Health Statistics, 1974.

All four surveys used a complex, multistage probability sampling design to provide representative samples of the civilian noninstitutionalized population of the United States. Means and standard deviations of the variables were weighted, using the weights supplied by the National Center for Health Statistics; these take into account the probability of selection, nonresponse, and poststratification adjustments. Multivariate regressions were estimated using the SAS regression program.

Table 1. Descriptive Data Concerning the Longitudinal Cycle 2 and 3 National Health Examination Survey Sample

Variable Measured	Cycle 2 Mean (SD)	Cycle 3 Mean (SD)	Correlation of Cycle 2 and Cycle 3 Values
Hours of television viewed per day	2.08 (1.10)	2.97 (1.39)	.28
WISC	101.57 (14.57)	102.12 (14.74)	.82
WRAT-Arithmetic	51.34 (9.46)	50.87 (9.69)	.70
WRAT-Reading	51.26 (9.55)	51.14 (9.75)	.86
Mother's years of schooling	11.04 (2.79)		
Father's years of schooling	11.00 (3.57)		
Race (Nonwhite = 1; white = 0)	.10 (.30)		
Birth order of child	1.99 (1.00)		
Family size	3.64 (1.90)		

NOTES: Cycle 2 was completed during 1963–65, and Cycle 3 was completed during 1966–70. N = 1,745.

Results

As presented in table 1, there was a fair amount of stability in measured IQ and the WRAT-A and WRAT-R scores over the four-year period between Cycle 2 and Cycle 3 (r = .82, .70, .86). These figures compare with WISC reliabilities of .92 to .95 on Wechsler's (1949) standardization sample, using the split-half technique, and of .77 using a test-retest technique over 49 months (Gehman and Matyas, 1956). In contrast, the measures of number of hours of television viewed only correlated .28 over the two time-points, presumably reflecting actual change in behavior, the fact that different individuals reported the data (parents at Cycle 2 and students at Cycle 3), as well as lack of reliability in the measures. As indicated in table 1, there was a substantial change in television viewing reported between Cycle 2 and Cycle 3, with the mean hours viewed increasing by an estimated 47%, from 2.1 to 3.0 hours per day.

Table 2. Coefficient Estimates from Different Models Relating Current Hours of Television Watched to IQ, WRAT-Arithmetic, and WRAT-Reading Test Scores

Variable	IQ	($p<$)	WRAT-A	($p<$)	WRAT-R	($p<$)
(A) Current TV	−2.23	(.0001)	−.97	(.0001)	−1.32	(.0001)
(B) Prior test score + (A)	−.51	(.0006)	−.25	(.05)	−.09	(.30)
(C) Season, region, population density + (B)	−.51	(.0006)	−.22	(.08)	−.10	(.26)
(D) All other + (C)[a]	−.25	(.10)	−.04	(.74)	−.003	(.98)
Final R^2 (adjusted)	.69		.51		.75	

NOTES: Children examined in Cycle 2 and Cycle 3 National Health Examination Survey. Each succeeding model incorporates the controls of the preceding models. N = 1,745.

[a] All other variables included mother's and father's education and age, income, number of children in family, child's birth order, race, and having conditions restricting activity.

Among children aged 12–17 in the sample, there was a substantial negative correlation between the amount of television viewed per day and measured IQ ($r = -.18$), the WRAT-A ($r = -.14$), and the WRAT-R ($r = -.17$) ($p < .0001$). These results closely parallel those of Gaddy (1986). The linear regression coefficients representing these relationships are presented in table 2. These coefficient estimates indicate that measured IQ in Cycle 3 is lower by 2.2 points for each additional hour of television viewed daily during the same Cycle, WRAT-A scores are lower by 1.0, and WRAT-R scores are lower by 1.3 points. We checked for nonlinearities in the relationship between television viewing and IQ, and there were none.

These bivariate regression results, however, do not take into account important other variables that could be causally responsible for producing variation in both the amount of television viewed and the measured scores. We first controlled for prior test scores, which substantially reduced the television/test score relationships (see row 2 of table 2). The considerable attenuation in the coefficients indicates that there was a substantial selection process taking place whereby children with lower scores during childhood tended to watch more television

Table 3. Coefficient Estimates of Different Models Relating Previous Hours of Television Watched to IQ, WRAT-Arithmetic, and WRAT-Reading Test Scores

Variable	IQ	(*p*<)	WRAT-A	(*p*<)	WRAT-R	(*p*<)
(A) Prior TV	−.90	(.0052)	−.46	(.0344)	−.71	(.0011)
(B) Prior test score + (A)	−.11	(.56)	−.05	(.76)	−.11	(.30)
(C) Season, region, population density + (B)	−.12	(.51)	−.06	(.70)	−.13	(.25)
(D) All other + (C)[a]	.11	(.55)	.14	(.37)	−.01	(.90)
Final R^2 (adjusted)	.70		.53		.77	

NOTES: Children examined in Cycle 2 and Cycle 3 National Health Examination Survey. Each succeeding model incorporates the controls of the preceding models. N = 1,745.

[a] All other variables included mother's and father's education and age, income, number of children in family, child's birth order, race, and having conditions restricting activity.

during adolescence, thereby accounting for a substantial part of the observed cross-sectional association between television viewing and test scores during adolescence.

When the remaining control variables, which included region of the country, size of place, seasonal variables, and the parent's socioeconomic status, were then introduced, the coefficients relating television viewing to test scores became statistically insignificant, just as in Gaddy's (1986) study. All of these control variables were measured at Cycle 2, and thus also indicate that a significant selection process was responsible for the observed negative uncontrolled relationship between test scores and television viewing.

It could be argued that while current television viewing might not have much effect on mental abilities, perhaps earlier viewing would. Therefore, we reestimated the same linear regressions cited in table 2, except we used the cross-lagged approach, regressing current mental abilities on the measure of television viewing several years earlier. These results, shown in table 3, closely parallel those in table 2. When television viewing is the only variable examined, there appears to be a modest but significant negative effect on mental abilities. But when other controls are added, the apparent effects become negligible and nonsignificant.

We also tested for a wide variety of possible interaction effects between television viewing and, for example, low mother's and father's education (< 12 years), high mother's and father's education (> 15 years), low income (< $4,000 in 1963–65), and race. None of these interactions had effects statistically significant at the $p < .05$ level. This suggests that the lack of relationship between television viewing and mental abilities holds for all socioeconomic strata in this study. In other words, we find no evidence for the suggestion that increased television viewing may help the relatively disadvantaged even if it hurts the relatively advantaged. Similarly, we entered the subject's age and the interaction of age with television viewing into the regressions to see whether there would be a differential effect of television viewing depending on age. These variables also were not significant.

In contrast to this evidence for no causal relationship between television viewing and test scores, other estimated coefficients in these regressions (see table 4) demonstrated substantial relationships between measured IQ at Cycle 3 and variables such as father's education, size of family, and birth order, after controlling for prior test score and all of the other variables indicated in table 2. In the case of WRAT-R, these were statistically significant independent relationships of father's education. Variables that independently predicted WRAT-A at Cycle 3 included mother's education and father's education.

Discussion

These data indicate that the amount of television viewing during childhood and adolescence is not a significant influence upon later test scores. But, these results are limited in a number of senses.

First, it might be argued that the sample employed is now out of date; however, television use was increasing during the period of study, and the timeframe of the data employed in these analyses was right in the middle of the long decline in Scholastic Aptitude Test Scores that so alarmed the nation (e.g., Time, 1975).

The analyses indicated no substantial influence during the later childhood years; however, it could be argued that a substantial portion of the cross-sectional relationship between the amount of television viewed and test scores emerged before these children were 6- to 11-year-olds. Because we do not have longitudinal data that predate the Cycle 2 time-point, this alternative explanation cannot be refuted. However, the correlations between amount of television viewed

Table 4. Estimated Regression Coefficients from Equations Predicting WISC, WRAT-R, and WRAT-A Scores, Cycle 2 and Cycle 3 of the NHES, 1963–1970

	Dependent Variables		
Independent Variable	WISC	WRAT-R	WRAT-A
Intercept	16.16***	2.77	8.12**
Television hours (Cycle 3)	.11	−.01	.14
Prior dependent measure (WISC, WRAT-R, or WRAT-A)	.76***	.84***	.65***
Central city (vs. major central city)	1.29*	.71	.96
Suburb	.44	.05	1.62**
Rural	.27	.08	1.16*
Winter (vs. fall)	−2.00**	.06	−.94
Spring	−.47	.45	.14
Summer	−.86	.72	−.36
Northeast (vs. West)	2.60***	.81*	−.38
Midwest	.84	.12	−.18
South	−1.38	−.55	−.36
Mother's education	.12	.10	.31**
Father's education	.32**	.12*	.16*
Income (missing = mean)	.23	.10	.16
Missing income (= 1)	−1.48	.40	−1.10
Mother's age (missing = mean)	−.08	.37	.50
Father's age (missing = mean)	.60	.05	.03
Number of children in family	.41**	.05	.24
Birth order	−.93**	−.06	−.15
Black (= 1)	−1.65*	−1.60**	−3.74***
Restricted physical activity	−.87	−1.59*	−1.12
R^2	.70	.53	.77

* $p < .05$. ** $p < .01$. *** $p < .001$.

among children aged 6–11 and test scores at the same age were only −.08 for all three tests.

Furthermore, the measure of television viewing used—the reported amount watched on a typical day—contains no information concerning program content, and the measure does not specify what other activities were taking place while television was being watched. It could be that the content of particular shows has positive effects, the content of other shows has negative effects, etc. We have no way of testing these hypotheses.

In addition, these findings are only applicable to individual differences. This was not a comparison of a group having television to one without it, but rather a study focusing on individual differences in the amount of television watched.

Finally, the single-item self-report measure of television viewing should exhibit a fair amount of unreliability and lack of validity, resulting in attenuated relationships observed between this variable and the outcome variables (Nunnally, 1982). One study of families with 5-year-olds indicated that parental global estimates of childhood viewing during a week are substantially correlated with parental estimates of viewing obtained via ten-day daily diaries (r = .48). These diary estimates have been shown to be strongly related to measures obtained using time-lapse videotapes of children watching television (Anderson et al., 1985). Ritchie, Price, and Roberts (1987) found high stability (around 0.8) in self-reported television viewing among children over a three-year period. These data indicate that the single global measure is reasonably reliable and valid. Even a substantial correction for attenuation would not result in a substantively meaningful relationship.

In spite of these limitations, our analyses indicate no significant causal relationship between the amount of television viewed during late childhood and adolescent test scores. The substantial cross-sectional relationships observed between these variables appear to be accounted for by the differential selection of children into patterns of television viewing. The panel data enables us to control for this selection process, and to obtain a more accurate estimate of causal effect, thus replicating and extending Gaddy's (1986) work to a younger population over a longer span of time.

References

Anderson, D., D. Field, P. Collins, E. Lorch, and J. Nathan (1985)
"Estimates of young children's time with television: A methodological comparison of parent reports with time-lapse video home observation." Child Development 56:1345–57.

Dietz, W. H., and S. L. Gortmaker (1985)
"Do we fatten our children at the television set? Obesity and television viewing in children and adolescents." Pediatrics 75:807–12.

Gaddy, G. D. (1986)
"Television's impact on high school achievement." Public Opinion Quarterly 50:340–59.

Gehman, I. H., and R. P. Matyas (1956)
"Stability of the WISC and Binet tests." Journal of Consulting Psychology 20:150–52.

Greenstein, J. (1954)
"Effects of television upon elementary school grades." Journal of Educational Research 48(3): 161–76.

Himmelweit, H. T., A. N. Oppenheim, and P. Vince (1958)
Television and the Child. London: Oxford University Press.

Hornik, R. C. (1978)
"Television access and the slowing of cognitive growth." American Educational Research Journal 15:1–15.
McLeod, J. M., M. A. Fitzpatrick, C. J. Glynn, and S. F. Fallis (1982)
"Television and social relations: Family influences and consequences for interpersonal behavior." Pp. 272–86 in P. Pearl, L. Bouthilet, and J. Lazar (eds.), Television and Behavior: Ten Years of Scientific Progress and Implications for the Eighties. Vol. 2. Washington, DC: U.S. Government Printing Office.
National Center for Health Statistics (1966)
Evaluation of Psychological Measures Used in the Health Examination Survey of Children Ages 6–11. Vital and Health Statistics. PHS Publication 1000, ser. 2, no. 15. Public Health Service. Washington, DC: U.S. Government Printing Office.
——— (1973)
Sample Design and Estimation Procedures for a National Health Examination Survey of Children. Rockville, MD: Health Resource Administration, U.S. Department of Health and Human Services, pub. no. (HRA) 74-1005.
——— (1974)
Plan and Operation of a Health Examination Survey of U.S. Youths 12–17 Years of Age. Rockville, MD: Health Resources Administration, U.S. Department of Health and Human Services, pub. no. (HRA) 75-1018.
Neuman, S. B. (1988)
"The displacement effect: Assessing the relation between television viewing and reading performance." Reading Research Quarterly 23:414–40.
Nunnally, J. (1982)
Psychometric Theory. 2nd Ed. New York: McGraw-Hill.
Ritchie, D., V. Price, and D. F. Roberts (1987)
"Television, reading, and reading achievement: A reappraisal." Communication Research 14:292–315.
Rubinstein, E. A. (1983)
"Television and behavior: Research conclusions of the 1982 NIMH report and their policy implications." American Psychologist 38:820–25.
Schramm, W., J. Lyle, and F. Parker (1961)
Television in the Lives of Our Children. Stanford, CA: Stanford University Press.
Singer, J. L. (1980)
"The power and limitations of television: A cognitive-affective analysis." In P. Tannenbaum (ed.), The Entertainment Function of Television. Hillsdale, NJ: Erlbaum.
Thompson, G. (1964)
"Children's acceptance of television advertising and the relation of televiewing to school achievement." Journal of Educational Research 58(4): 171–75.
Time (1975)
"Can't anyone here speak English?" August 25, 34–36.
Wechsler, D. (1949)
Wechsler Intelligence Scale for Children. New York: Psychological Corporation.
Williams, P. A., E. H. Haestel, G. D. Haertel, and H. J. Walberg (1982)
"The impact of leisure-time television on school learning: A research synthesis." American Educational Research Journal 19:19–50.
Witty, P. (1967)
"Children of the TV era." Elementary English 44:528–35, 554.
World Almanac (1986)
New York: Newspaper Enterprise Association.

NOTE

ONLY THE FOLLOWING PAGES OF THIS

PUBLICATION HAVE BEEN FILMED

Cov. 1, 2, Contents Pages:

1, 2, Pgs: 24-49, 80-127,

135-148

AGG10-1

Public Opinion Quarterly

Published for the American
Association for Public Opinion
Research by the University of
Chicago Press

NORBERT SCHWARZ, FRITZ STRACK, and HANS-PETER MAI Assimilation and Contrast Effects

NICHOLAS DANIGELIS and STEPHEN CUTLER The Aging-Conservatism Hypothesis

SCOT BURTON and EDWARD BLAIR Behavioral Frequency Questions

DANIEL LINZ and others Estimating Community Standards

ERIC PLUTZER and MICHAEL MCBURNETT Family Life and American Politics

IN MEMORIAM Robert T. Bower

THE POLLS Review Reply: Running the R.A.C.E.? *by Lawrence Bobo*

Trends: Attitudes toward the USSR *by Alvin Richman*

COMMENT Comment *by Charlotte Steeh*
Response *by Jack Citrin, Donald Green, and David Sears*

BOOKS Review Essay *by Richard T. Campbell*

8

Spring 1991

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR Advisory Committee

Eleanor Singer, Chair (1987–91)
Herbert Abelson (1988–92) Gladys Engel Lang (1989–93)
W. Russell Neuman (1990–94)

Stanley Presser and Kathleen Frankovic,
AAPOR Councillors-at-Large

Howard Schuman, *Editor*

Manuscript Editor	*Office Manager*	*Assistant to the Editor*
John D. Vander Weg	Carlotta Nosel	Katherine Bischoping

The Polls
Reviews: Stanley Presser, Seymour Sudman
Trends: Robert Y. Shapiro

EDITORIAL BOARD

Herbert I. Abelson
Robert P. Abelson
Sandra J. Ball-Rokeach
Larry M. Bartels
Nathaniel Beck
Sandra H. Berry
Lawrence D. Bobo
Pamela Conover
Jean M. Converse
Stanley Feldman
Richard J. Fox
Deborah R. Hensler
Shanto Iyengar
Graham Kalton
Gary King
Jon A. Krosnick
Richard A. Kulka
Kurt Lang
McKee J. McClendon
William J. McGuire
Michael MacKuen
Jack M. McLeod
Elizabeth Martin
John Mueller
Harold Nisselson
Helmut Norpoth
Vincent Price
Willard L. Rodgers
Norbert Schwarz
Robert Y. Shapiro
Lee Sigelman
Tom W. Smith
John Sullivan
John Zaller

Public Opinion Quarterly (ISSN 0033-362X) Volume #55, 4 issues (1991)

© 1991 by the American Association for Public Opinion Research. Published quarterly: Spring, Summer, Autumn, and Winter, by The University of Chicago Press, 5720 S. Woodlawn Avenue, Chicago, Illinois 60637. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $40.00; individuals 1 year $20.00. WAPOR member subscription rate, U.S.A.: individuals 1 year $18.00. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $18.00. Other countries add $3.00 for each year's subscription to cover postage.

Japanese subscription agent: Kinokuniya Company Ltd.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to: Journals Fulfillment Department, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Institute for Social Research
P.O. Box 1248, Ann Arbor, MI 48106
Business, Subscriptions, Back Issues: The University of Chicago Press,
5801 South Ellis, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, New Jersey 08540

Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. ♾

PUBLIC OPINION QUARTERLY

VOLUME 55 / SPRING 1991 / NUMBER 1

Editorial 1

Articles

Assimilation and Contrast Effects in Part-Whole Question Sequences: A Conversational Logic Analysis 3
NORBERT SCHWARZ, FRITZ STRACK, and HANS-PETER MAI

Cohort Trends in Attitudes about Law and Order: Who's Leading the Conservative Wave? 24
NICHOLAS L. DANIGELIS and STEPHEN J. CUTLER

Task Conditions, Response Formulation Processes, and Response Accuracy for Behavioral Frequency Questions in Surveys 50
SCOT BURTON and EDWARD BLAIR

Estimating Community Standards: The Use of Social Science Evidence in an Obscenity Prosecution 80
DANIEL LINZ, EDWARD DONNERSTEIN, KENNETH C. LAND, PATRICIA L. MCCALL, JOSEPH SCOTT, BRADLEY J. SHAFER, LEE J. KLEIN, and LARRY LANCE

Family Life and American Politics: The "Marriage Gap" Reconsidered 113
ERIC PLUTZER and MICHAEL MCBURNETT

In Memoriam

Robert T. Bower, 1919–1990 128

The Polls

Review Reply: Running the R.A.C.E.? 131
LAWRENCE BOBO

Trends: Changing American Attitudes toward the Soviet Union 135
ALVIN RICHMAN

Comment

Comment 149
CHARLOTTE STEEH

Reply 150
JACK CITRIN, DONALD GREEN, and DAVID SEARS

Book Review Essay

The Sage University Papers on Regression Analysis 151
RICHARD T. CAMPBELL

LIST OF MANUSCRIPT REFEREES, 1989–1990 156

NOTICE TO CONTRIBUTORS 158

COHORT TRENDS IN ATTITUDES ABOUT LAW AND ORDER: WHO'S LEADING THE CONSERVATIVE WAVE?

NICHOLAS L. DANIGELIS and
STEPHEN J. CUTLER

Abstract Aging is commonly assumed to be associated with the adoption of increasingly conservative social and political attitudes. Empirical tests of this assumption have provided little support, but most analyses have focused on attitude domains where public opinion was shifting in a liberal direction. Using three items about law and order, a domain marked by conservative trends, this paper evaluates the aging-conservatism hypothesis with data drawn from 18 national surveys covering the period 1959–85. Attitude changes in four cohorts are analyzed with Taylor's hierarchical goodness-of-fit procedures for evaluating trends in public opinion. The findings show that each of the cohorts has participated in the shift toward more conservative opinions on law and order issues, and that either constant or zero differences characterize the trends in percentage differences between the oldest and the other cohorts. In the absence of differential rates of change, we conclude that older cohorts are no more likely than younger cohorts to adopt conservative law and order attitudes and that period effects are influencing each of the cohorts equally.

Introduction

THE ARGUMENTS

A widely-held assumption is that growing old is accompanied by the adoption of increasingly conservative social and political attitudes. For

NICHOLAS L. DANIGELIS is Associate Professor of Sociology at the University of Vermont. STEPHEN J. CUTLER is Professor of Sociology and the Bishop Robert F. Joyce Distinguished University Professor of Gerontology at the University of Vermont. This is a revised version of a paper presented at the annual meeting of the Gerontological Society of America, Washington, DC, November, 1987. The research was supported by a grant from the National Institute on Aging (R01 AG03420). The data used in the analysis were obtained from the Inter-University Consortium for Political and Social Research and the Roper Center. The authors are especially appreciative of the painstakingly careful work of our research assistant, Stephanie Robert, and our administrative assistant, Sarah Gilmore.

Public Opinion Quarterly 55:24–49 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-01/$2.50

instance, commenting on possible consequences of the older age structures characteristic of zero–population growth societies, Coale (1964, 57) suggests that "such a population would be cautious, conservative, and full of regard for the past." Similarly, the Commission on Population Growth and the American Future (1972, 96) noted that "one concern often expressed about an older age structure is that there will be a larger proportion of the population who are less adaptable to political and social change, thus suggesting the possibility of 'social stagnation.' " As a last example, Berelson and Steiner (1964, 573; emphasis added), in commenting on age differences in tolerance of ideological nonconformity reported by Stouffer (1955), state that ". . . *age appears to bring with it not only a developing conservatism*, . . . but also a higher degree of authoritarianism." They go on to suggest that "as each new generation in this country is better educated than the one before, it is more tolerant; but *as it grows older it becomes less so*."

Aging at the later stages of the life course, then, is commonly seen to be associated with increasing conservatism, that is, with resistance or opposition to change in general and to change that would benefit disadvantaged segments of the population; with a higher valuation of social order; with an increasing emphasis on authority and obedience; and with the adoption of a generally restrictive, rather than permissive and tolerant, attitude toward human behavior (Glenn 1974).

While theoretical support for this assumption about the course of attitude change with aging takes several forms (Cutler 1977; Glenn 1980; Sears 1981), the linkage between growing old and conservatism has come under criticism in terms of its conceptual (Campbell 1971), empirical (Campbell and Strate 1981; Cutler and Kaufman 1975; Danigelis and Cutler 1989; Herzog 1979; Smith 1981), and methodological (Riley 1973; Schaie 1967) bases. Perhaps the most compelling evidence casting doubt on the aging-conservatism argument has come from studies that have examined changes in social and political attitudes during periods of growing liberalism in public opinion (e.g., Danigelis and Cutler 1989; Smith 1981; Thornton, Alwin, and Camburn 1983). This research consistently finds the attitude change occurring among older cohorts to be in the same liberal direction, and often at the same rate, as that which characterizes younger cohorts. What is missing, however, and what would provide an important complementary test of the aging-conservatism thesis, is an examination of attitudes in a domain where a clear conservative trend in public opinion has been occurring.

The resurgence of conservative attitudes in the United States within the last several years has been asserted with respect to a number of social, economic, and political issues, but perhaps nowhere is the trend away from the liberalism of the 1960s more apparent than in the area of attitudes about law and law enforcement (Exter and Barber 1986; see also Smith 1985). If any set of issues is likely to support the idea that trends toward conservatism occur earlier and are especially pronounced among the aged, it should be in the area of law and law enforcement attitudes. In general, if the aging-conservatism argument is correct, a trend comparison of the attitudes of different birth cohorts should show older cohorts becoming more conservative sooner, and at a faster rate, than younger cohorts. The purpose of this paper is to provide a direct test of this hypothesis.

THE EVIDENCE

Overall Trends. Evidence on attitudes toward gun control, the courts, and capital punishment is particularly instructive. Trend data presented for the last 20 to 30 years in Stinchcombe et al. (1980, chap. 2), for example, show that Americans have become somewhat more opposed to gun control, more critical of the courts, and more supportive of capital punishment.

There are important caveats about the strength of these trends, however. Concerning gun control, the trend away from liberalism appears slight, at best (Stinchcombe et al. 1980, 36). With respect to capital punishment, the recent trend toward conservatism may be misleading, because data extending back to the 1930s (Stinchcombe et al. 1980, 28) show the trend to be curvilinear: a swing away from support for capital punishment between the 1930s and the 1960s followed by a swing back toward support over the last 20 years or so. Only the public's attitude toward the courts appears to show a definite trend away from liberalism, but here, as with the question on gun control, data go back only to the mid-1960s (Stinchcombe et al. 1980, 31).

Aging and Attitudes Toward Law and Order Issues. Evidence concerning the association between aging and attitudes toward gun control, the courts, and capital punishment comes from two sources. The first is indirect and based on the presumed association between attitudes that are tied to both age and these law and order questions. Deemed to be of particular importance is "fear of crime." For one thing, it is argued that older people, regardless of how little they may, in fact, be victimized, are significantly more likely than younger people to be afraid of crime (Alston 1986; Lawton et al. 1976; Yin 1985). Furthermore, people who fear crime are alleged to be much more likely to believe punishment can be effective in fighting crime and therefore support more stringent approaches to crime control (DeFronzo 1979; Thomas and Foster 1976).[1]

---

[1] A similar argument can be made for "feelings of retribution" as an intervening variable (see Warr and Stafford 1984).

The available evidence, however, indicates that the association between fear of crime and attitudes about capital punishment is negligible (Taylor, Scheppele, and Stinchcombe 1979), exists for blacks but not for whites (Combs and Comer 1982), and varies by gender (Rankin 1979). Additionally, when one examines attitudes toward law and order issues over time, any association between fear and these attitudes may be more properly attributed to increases in the violent crime rate (Rankin 1979). As Stinchcombe et al. (1980, 13) argue, "The increase in fear is not large enough to explain the movement of punitiveness [toward the conservative position] over time."

Direct evidence concerning the association between aging and attitudes toward law and law enforcement issues is somewhat scanty. Cross-sectional data are contradictory. Alston (1976), Sengstock and Liang (1979), Vidmar and Ellsworth (1976), and Williams, Thomas, and Singh (1983) assert that older people support harsher sanctions for criminals, including the use of capital punishment. Thomas and Howard (1978), however, argue that age differences are small, while Ruggiero and Weston (1977) find older people to be *less* punitive regarding the death penalty when gender is controlled.

More pertinent evidence comes from Exter and Barber (1986, 32–33), who found the oldest respondents to be the most conservative and the youngest the most liberal. However, "[t]he gap between the age groups most likely and least likely to think the courts are too lenient shrank from 14 percentage points in 1975 to 8 points in 1985," and "[t]he gap between the age group most supportive of capital punishment and the one least supportive declined from 13 percentage points to 4." For gun control, the negligible initial differences between oldest and youngest persisted.

A more direct test of the aging-conservatism notion is provided by a simple cohort analysis based on the data presented by Exter and Barber. If one compares the change in attitudes on gun control, the courts, and capital punishment among those who were in their 20s and 50s in 1975 with those who were in their 30s and 60s in 1985, each of the two groups shifted in a conservative direction. However, the younger group has become conservative to a greater degree than the older one on items about the courts and capital punishment, while change in attitudes on gun control was the same for each cohort.

Altogether, the cross-sectional and longitudinal data available do not provide clear support for the aging-conservatism argument nor do they clearly refute it. In the remainder of the paper, we provide a direct test of the idea that the adoption of increasingly conservative views about law and order issues is especially pronounced with aging among persons at the later stages of the life course.

Table 1. Law and Order Attitudes Used

Gun Control	
Question	Would you favor or oppose a law which would require a person to obtain a police permit before he or she could buy a gun?
Mnemonic	GUNLAW
Liberal response	Favor
Surveys used	AIPO #616, 1959; AIPO #681, 1963; AIPO #704, 1965; AIPO #749, 1967; NORC GSS 1972–77, 1980, 1982, 1984, 1985
Court System	
Question	In general, do you think the courts in this area deal too harshly or not harshly enough with criminals?
Mnemonic	COURTS
Liberal responses	Too harshly, about right (volunteered)
Surveys used	AIPO #716, 1965; AIPO #773, 1969; NORC GSS 1972–78, 1980, 1983–85.
Capital Punishment	
Question	Do you favor or oppose the death penalty for persons convicted of murder?
Mnemonic	CAPPUN
Liberal response	Oppose
Surveys used	AIPO #625, 1960; AIPO #704, 1965; AIPO #774, 1969; NORC GSS, 1972–78, 1980, 1982–85.

Methods

In order to test the idea that aging is associated with the adoption of an increasingly conservative view toward law and order, we employ data from 18 national surveys to analyze trends in three replicated items spanning the period 1959–85. The analysis is based on a cross-sequential design (Schaie 1983) which will allow us to follow any attitude change of different age-cohorts during this time. Below we describe the data sets and variables employed, explain the particular statistical analysis selected, and state the specific research hypotheses to be tested.

DATA SET SELECTION AND DEPENDENT VARIABLES

Table 1 shows the three attitude questions whose patterns of change will be examined, the exact wording of the questions, their mnemonic names, the category or categories of the "liberal" response, and the surveys from which they have been taken. The question on gun control covers the longest period, 26 years (1959–85), while the item on the courts covers the shortest, 20 years (1965–85).[2]

Respondents who have not provided substantive answers to the questions have been omitted from the analysis. This includes the missing data categories "don't know," "no opinion," "refused," and "no answer." Also missing are nonwhite respondents, almost all of whom are black. Our analysis is limited to whites for two reasons. First, another part of the larger study on which the analysis described here is based focuses on attitudes about race relations, the questions for which were not asked of blacks in the earlier surveys. Second, however, there is some evidence to believe that the attitudes of blacks on law and order issues may not follow the same patterns as those of whites (Combs and Comer 1982; Stinchcombe et al. 1980; Taylor, Scheppele, and Stinchcombe 1979). Given the evidence on the relative liberalism of blacks regarding law and order issues (Stinchcombe et al. 1980, 125), it is quite likely that there will be greater changes among white cohorts than among comparable black ones. Because the black samples in the surveys used are too small to divide into cohorts in the same way as is done with whites, we decided to omit all nonwhite respondents.

Several criteria governed the choice of the questions and surveys described in table 1. First, it was decided at the outset to include only nationally representative surveys. Although data sets containing replicated items exist for smaller area samples (e.g., the Detroit Area Studies), cohort analyses based on such surveys are extremely likely to introduce unknown effects attributable to migration among an open population (Hyman 1972).

Second, an effort was made to select surveys covering a sufficient span of time so that trends in attitude change would be less likely to reflect short-term fluctuations. As indicated above, the time periods covered extend between 20 years (COURTS) and 26 years (GUNLAW).

A third requirement for the type of analysis to be conducted is that the items be exactly or nearly exactly replicated. This assures that any observed changes in attitudes are due to reasons other than substantive changes in the wording of the questions. While full comparability is preferred (and is the case with two of the items), minor changes in question wording—such as the formal balancing of response alternatives (Schuman and Presser 1981)—can often be tolerated. For instance, in regard to the question on capital punishment (CAPPUN), the AIPO surveys and the 1972 and 1973 General Social Surveys asked "Are you in favor of the death penalty for persons convicted of murder?" while the more recent General Social Surveys asked "Do you favor or oppose the death penalty for persons convicted of murder?" Although the explicit inclusion of opposition as a response alternative might be expected to legitimate and increase the proportion giving this response, the percentage of respondents favoring the death penalty actually increased from 60% to 63% between 1973 and 1974 when the new wording was first used.

---

[2] In the list of surveys, "AIPO" designates Gallup Polls conducted for the American Institute for Public Opinion. These surveys were obtained from the Roper Center. "NORC GSS" refers to the National Opinion Research Center's General Social Surveys.

INDEPENDENT VARIABLES

In addition to the three dependent variables, four independent variables have been selected from each of the surveys. *Time of measurement* is indicated by the last two digits of the year in which the surveys were conducted. *Age of respondent* allows construction of a *cohort* measure that distinguishes four birth cohorts. The earliest year for any of the dependent variables (1959) suggests the following distinctions: respondents born between 1930 and 1938, who were between 21 and 29 years old in 1959 ("youngest"); those born between 1920 and 1929, who were between 30 and 39 years old in 1959 ("middle young"); those born between 1910 and 1919, who were between 40 and 49 years old in 1959 ("middle old"); and those born before 1910, who were 50 years old or older in 1959 ("oldest"). Respondents born after 1938 were dropped. *Gender* obviously is a dichotomy: males and females. *Level of educational attainment* has been divided into three groups: respondents who did not complete high school, those completing 12 years of education or high school graduates, and those who have had one or more years of college education.[3]

We employ gender and education in the analysis for two reasons. First, early surveys, especially those conducted during the 1930s and 1940s, used sampling designs and procedures that tended to underrepresent certain segments of the population: females, southerners, blacks, persons with low levels of education, and nonvoters. For the present investigation, this problem is less consequential. The earliest survey we will be using was conducted in 1959, and "by the mid-1950s, all of the major survey organizations were designing their usual samples to represent the total adult noninstitutionalized civilian population and had changed from quota control to 'modified probability' samples" (Glenn 1975, 10–11).

---

[3] Education has been trichotomized for three reasons. First, the earlier surveys do not provide as much detail in distinguishing formal schooling levels as has been done in the more recent NORC General Social Surveys, necessitating the use of cutoff points common to all studies. Second, the three categories used represent important formal schooling achievement differences and, consequently, important life-style differences. Finally, the analytical technique employed is facilitated by a relatively small number of categories for each of the control variables.

While such changes in sampling design and more complete representation should minimize any problem of systematic underrepresentation in our earlier surveys, the potential effects of any residual underrepresentation in the 1950s should be further reduced by the use of gender and education to address a second issue: compositional changes due to mortality differentials. The characteristics of persons lost through mortality are not randomly distributed. Among the major sources of differential attrition are gender, race, and socioeconomic status. Over time, and especially among older cohorts, compositional changes occur, so that a cohort typically contains an increasing proportion of women, whites, and persons of higher socioeconomic status. The problem for the cohort analyst is that compositional changes can affect the dependent variable of interest if that variable is related to the dimensions along which differential attrition takes place. Thus, if differential attrition produces a cohort with an increasingly higher level of educational attainment, and if the dependent variable of interest is related to education, the compositional change can have the effect of artifactually altering the dependent variable score even if no attitude change is occurring.

These correlations require that "true" changes somehow be separated from attitude changes due to compositional effects. A viable approach to this problem is to examine cohort scores on the dependent variable after the effects of variables associated with differential attrition have been removed. Partialling out effects due to variables such as education and gender and then examining cohort scores on the dependent variable provides some assurance that any observed intracohort attitude changes are not attributable to the changing cohort composition.[4]

---

[4] Because the samples to be examined are restricted to whites, such a correction is not necessary for race in this analysis. Gender initially was thought to be an important control variable for another reason. Given the evidence regarding which groups fear crime most, it was believed that gender and age might interact statistically to produce different patterns of cohort change in attitudes among men and women. As Stinchcombe et al. (1980, 57) point out, fear of crime increases substantially with age for men but not for women. If one puts together the evidence on a possible gender/age interaction regarding fear with what we know about fear of crime making people more punitive and, thus, more likely to be conservative on law and crime issues (Stinchcombe et al. 1980, chap. 3), then the cohort trends ought to differ for women and men. However, after separately analyzing the attitude trends reported in table 4 for males and females, we found similar patterns of results. Therefore, we report analyses that assume only additive effects from the control variables.

ANALYSIS DESIGN, DATA ANALYSIS, AND RESEARCH HYPOTHESES

The overall data analysis approach is a cross-sequential design that allows intra- and intercohort comparisons over time (Schaie 1983). The specific analytic technique employed is what has come to be called "survey metric analysis" (Firebaugh and Davis 1988; Smith 1981), an extension of Davis's *d*-system analysis to trend data (Davis 1975; Taylor 1980). The combination of these approaches will enable us to model cohort changes on law and order questions, testing the idea that shifts to more conservative attitudes are especially pronounced among the older cohorts. However, because the cross-sequential design focuses on cohort and period (or time-of-measurement) effects, it does not provide a direct test for the existence of age effects within cohorts. Age and period effects are confounded in a cross-sequential design and cannot be easily separated unless simplifying (and often unrealistic) assumptions are made about the absence of one or the other effect. Following Riley 1973 and Glenn and Frisbie 1977, therefore, we will be more interested in describing the patterns of attitude change on these law and order items than in accounting for the precise sources of change.

Each of the three attitude items was dichotomized into "liberal" and "conservative" responses as described in table 1, then three types of dependent variables were considered. The first dependent variable is simply the percentage of the total sample (i.e., the four cohorts combined) giving the liberal response for each item: the percentage *in favor* of requiring gun permits, the percentage believing that the courts are doing *alright* or are dealing *too harshly* with criminals, and the percentage *opposing* capital punishment.[5] The second dependent variable is the liberal percentage for each cohort, standardized for gender and education differences among the cohorts (to control for changing cohort composition). This was accomplished by using the total sample's gender and education distribution for the first year in which a variable appeared as a standard and adjusting each liberal percentage value for each cohort accordingly (see Mueller, Schuessler, and Costner 1977, 132–46). The third dependent variable is the *adjusted difference* in liberal percentage between the oldest cohort and each younger cohort (younger − oldest). To calculate this adjusted difference, each younger cohort is compared with the oldest cohort separately for each year, with gender and education controlled.

Each dependent variable was then plotted over time and the results graphed. The total sample, cohort-specific, and cohort difference changes over time were evaluated statistically using procedures developed and described by Taylor (1980).[6] Through a hierarchical sequence of goodness-of-fit tests, the Taylor procedures determine whether change is occurring and, if so, the type of change pattern observed.

---

[5] We also considered combining each of the three questions into an index. However, as the survey years listed for each question in table 1 show, there are too few surveys, especially in the earlier years, in which all three questions were asked.

With respect to the first two dependent variables (total sample and cohort-specific trends in liberal percentage), the models tested are: a constant liberal percentage over time, an increasing liberal percentage over time, a decreasing liberal percentage over time, and a nonlinear change in liberal percentage over time. For the third dependent variable, intercohort comparisons of adjusted percent differences, we test five models: no difference over time in liberal percentage, a constant difference over time, an increasing difference over time, a decreasing difference over time, and a nonlinear change in differences over time.

Based on the evidence cited earlier, we expect a trend away from the liberal response for the total sample (all four cohorts combined) for all three questions during the period covered. This means that the first dependent variable analysis should show a *decreasing* percentage of liberal respondents over time. When adjustments are made for changing cohort composition, and cohort trends are examined separately and compared with one another, the aging-conservatism thesis argues that the oldest cohort should display a markedly greater tendency toward conservatism than the other cohorts. While the trends for all cohorts likely will be toward decreasing liberalism, as shown by a negative slope, the trend line of the oldest cohort should be steeper (and the siope larger) than those of the other cohorts (second dependent variable analysis). Therefore, when the adjusted differences in liberal percentage between the oldest and each younger cohort are examined (third dependent variable analysis), we should observe *increasing* differences over time.

Results

TOTAL SAMPLE TRENDS

Figure 1 shows the trends in liberal percentage for all four cohorts combined for each of the three questions. While more conservative

---

[6] We also followed Taylor's suggestion (1980, 90) to correct the variances of the percentages (or percentage differences) when the effective size of a sample is in question due to lowered efficiency in sampling procedures (e.g., the use of quota sampling in the early surveys). Because increasing the variances has the overall effect of increasing the likelihood of no differences between cohorts, we also ran the analyses without adjusting for sample efficiency differences. The results were quite similar to those reported in tables 2–4.

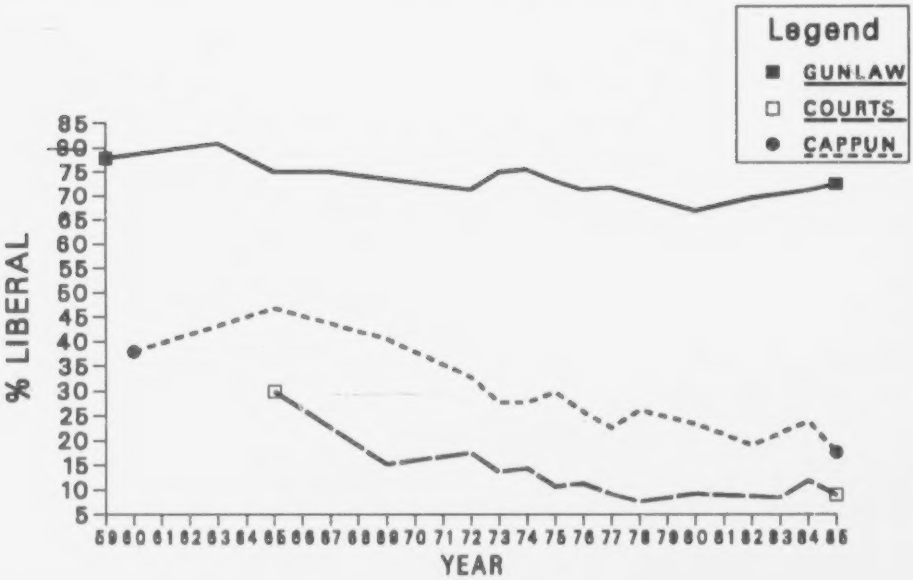


Figure 1. Total Sample Trends in Law and Order Attitudes by Type of Question and Year, 1959–85

attitudes for each cohort are evident in 1985 than in the first years, the three trend-lines are somewhat different. There is a slight overall reduction in liberal sentiment about gun control from 1959 to 1985, but much fluctuation in the intervening years. The question on courts produces a substantial drop in the liberal position between 1965 and 1969, followed by a less pronounced decline in the 1970s. Finally, sentiment against the death penalty initially increases (1960–65), but then attitudes quickly become more conservative as the percentage of liberal responses steadily decreases.

Table 2 summarizes the application of Taylor's (1980) extension of Davis's *d*-system analysis to trend data. The best fitting model for each item is one that has a substantial negative linear component, because each trend is closer to a negative linear trend model than to either a constant liberal percentage or a nonlinear trend model. As expected, therefore, attitudes toward crime and the law did tend to become less liberal, or more conservative, from 1959 through 1985. The slight differences in the magnitude of the slopes in table 2 are consistent with the separate trends shown in figure 1. The least change occurs for the gun control question ($b = -.37$), while the trends away from liberal attitudes are more dramatic for the courts ($b = -.76$) and capital punishment ($b = -1.04$) questions.

Table 2. Hierarchical Goodness-of-Fit Tests for Total Sample Trends

Model	χ^2	*df*	*p*	Decision
	Percentage Favoring Gun Permit (GUNLAW)			
a. Percent = constant	82.8	13	<.001	Reject
b. Percent = linear trend	29.1	12	<.001	Reject
c. Difference, *a* − *b*	53.7	1	<.001	Significant
Final: Substantial negative linear component	% = 101.0 − .37 year			
	Percentage Believing Courts Too Harsh, OK (COURTS)			
a. Percent = constant	240.2	12	<.001	Reject
b. Percent = linear trend	98.4	11	<.001	Reject
c. Difference, *a* − *b*	141.8	1	<.001	Significant
Final: Substantial negative linear component	% = 70.8 − .76 year			
	Percentage Opposing Capital Punishment (CAPPUN)			
a. Percent = constant	400.1	14	<.001	Reject
b. Percent = linear trend	82.8	13	<.001	Reject
c. Difference, *a* − *b*	317.3	1	<.001	Significant
Final: Substantial negative linear component	% = 106.9 − 1.04 year			

NOTE: Periods covered by the series are 1959–85 for GUNLAW, 1965–85 for COURTS, and 1960–85 for CAPPUN.

COHORT SPECIFIC TRENDS

The data in table 3 summarize statistical evaluations of each cohort trend, standardized for gender and education. Out of the twelve tests, nine show that cohort attitudes toward law and law enforcement issues have either clearly become more conservative (negative linear trend) or have tended toward conservatism (substantial negative linear component). Overall, the prediction that recent years have seen a decline in liberal attitudes on law and law enforcement issues is basically supported for each of the cohorts.

Table 3. Hierarchical Goodness-of-Fit Tests for Trends by Cohort, Adjusted for Sex and Education Composition Differences

	Cohort							
	Youngest (b. 1930–38)		Middle-Young (b. 1920–29)		Middle-Old (b. 1910–19)		Old (b. before 1910)	
Model	*p*	Decision	*p*	Decision	*p*	Decision	*p*	Decision
	Percentage Favoring Gun Permit (GUNLAW)							
a. Percentage = constant	>.05	Accept	<.01	Reject	<.05	Reject	>.05	Accept
b. Percentage = linear trend	—	—	>.05	Accept	>.05	Accept	—	—
c. Difference, *a* − *b*	—	—	<.001	Significant	<.02	Significant	—	—
Final	Constant % = 70.2		Negative linear trend % = 106.6 − .46 yr.		Negative linear trend % = 95.9 − .30 yr.		Constant % = 77.1	

	Percentage Believing Courts Too Harsh, OK (COURTS)							
a. Percentage = constant	<.001	Reject	<.001	Reject	<.001	Reject	<.001	Reject
b. Percentage = linear trend	<.05	Reject	<.01	Reject	<.001	Reject	<.001	Reject
c. Difference, *a* − *b*	<.001	Significant	<.001	Significant	>.05	n.s.	<.05	Significant
Final	Substantial negative linear component % = 86.4 − .96 yr.		Substantial negative linear component % = 72.8 − .78 yr.		Nonlinear trend % = 8.4		Substantial negative linear component % = 49.5 − .48 yr.	
	Percentage Opposing Capital Punishment (CAPPUN)							
a. Percentage = constant	<.001	Reject	<.001	Reject	<.001	Reject	<.001	Reject
b. Percentage = linear trend	>.05	Accept	>.05	Accept	<.05	Reject	<.001	Reject
c. Difference, *a* − *b*	—	—	—	—	<.001	Significant	<.001	Significant
Final	Negative linear trend % = 116.3 − 1.18 yr.		Negative linear trend % = 107.4 − 1.07 yr.		Substantial negative linear component % = 111.8 − 1.13 yr.		Substantial negative linear component % = 108.4 − 1.03 yr.	

NOTES: All results based on χ^2 tests with appropriate degrees of freedom (see Taylor 1980); n.s. = not significant. Periods covered by the series are 1959–85 for GUNLAW, 1965–85 for COURTS, and 1960–85 for CAPPUN.

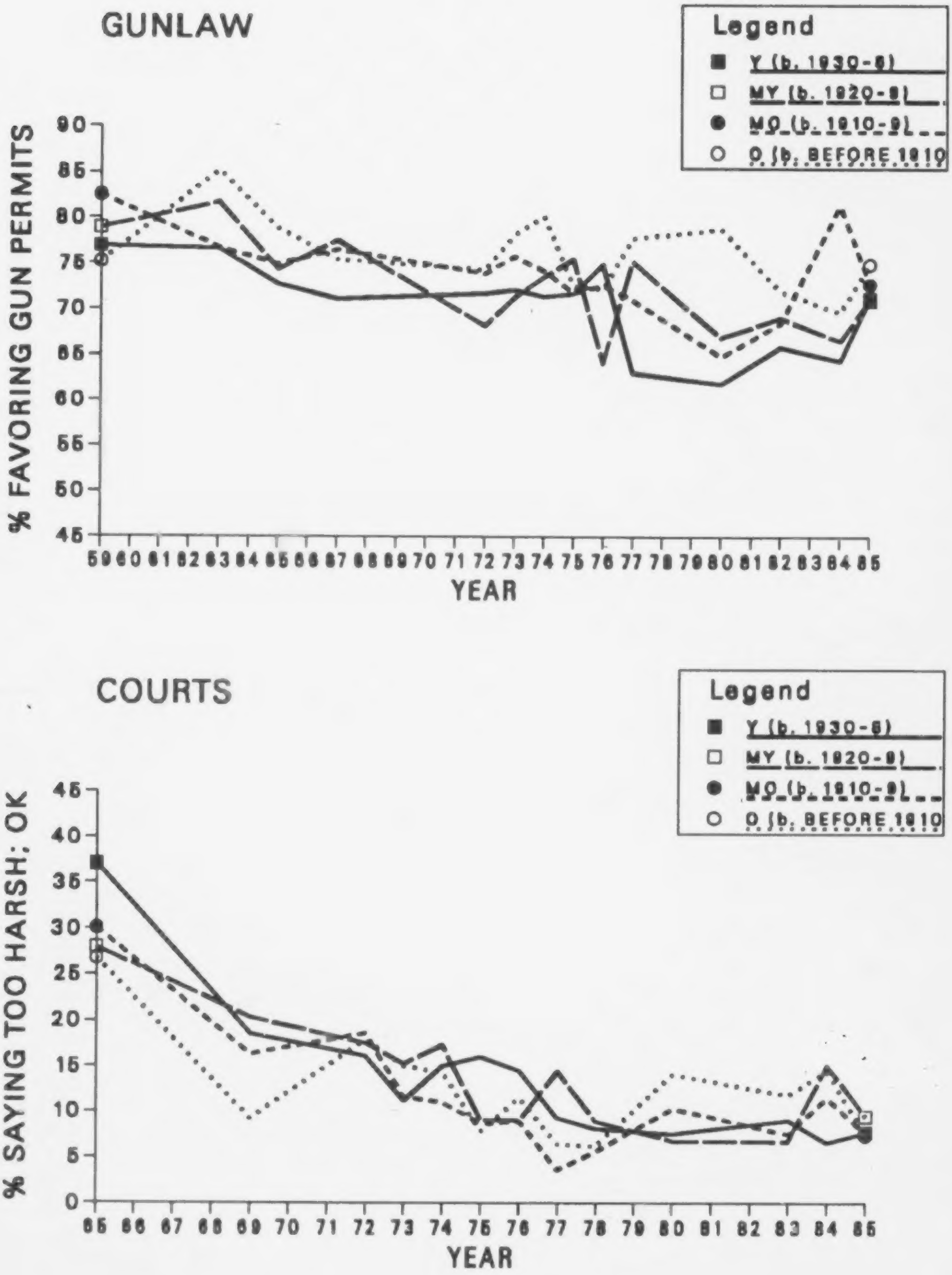


Figure 2. Trends in Law and Order Attitudes by Cohort and Year, Adjusted for Sex and Education Composition Differences (*top:* Gun Laws, *bottom:* Courts)

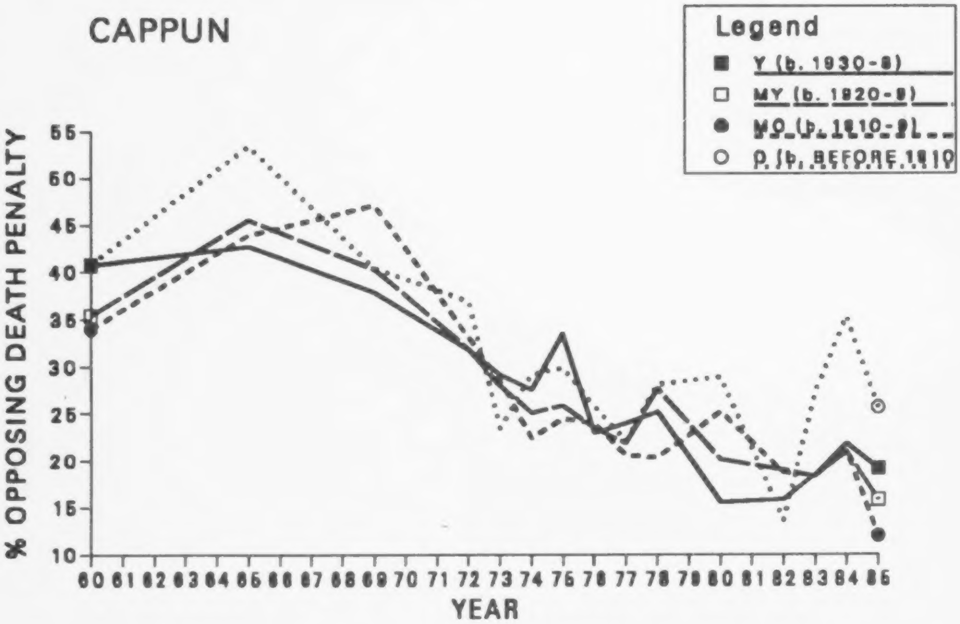


Figure 2 (*continued*). Trends in Law and Order Attitudes by Cohort and Year, Adjusted for Sex and Education Composition Differences (Capital Punishment)

When the three questions are examined separately, support for the prediction of decreasing liberalism strengthens as one reads down table 3 from the gun control, to the courts, to the capital punishment question. At one end, only the two middle cohorts seem to have become more conservative on gun control because the level of liberal attitudes among the youngest and oldest cohorts has tended to remain constant. At the other end, all four cohorts have clearly (or apparently) become more conservative on capital punishment. Furthermore, all four slopes appear to have approximately the same magnitude (around – 1.00).

A graphic portrayal underscores both the overall message and the item-by-item differences summarized by the data in table 3. While negative trends do tend to dominate figure 2, each of the four standardized cohort trends in gun control attitudes fluctuates up and down quite substantially. For the questions on the courts and capital punishment, the fluctuations are less dramatic. While the initial drop in liberal attitudes appears to be fairly steep only for the courts question, attitudes toward both the courts and capital punishment become markedly more conservative during the period 1965–72. The apparent discrepancy is due to the different starting dates (1965 and 1960, respectively) for each question.

Table 4. Hierarchical Goodness-of-Fit Tests for Trends in Cohort Differences, Adjusted for Sex and Education Composition Differences

	Cohort Comparison					
	Youngest–Oldest		Middle-Young–Oldest		Middle-Old–Oldest	
Model	*p*	Decision	*p*	Decision	*p*	Decision
	Percentage Favoring Gun Permit (GUNLAW)					
a. Percentage diff = 0	<.001	Reject	<.01	Reject	<.01	Reject
b. Percentage diff = constant	>.05	Accept	>.05	Accept	<.05	Reject
c. Difference, *a* − *b*	<.001	Significant	<.001	Significant	—	—
d. Percentage diff = linear trend	—	—	—	—	<.05	Reject
e. Difference, *b* − *d*	—	—	—	—	>.05	n.s.
Final	Constant % diff = −9.1		Constant % diff = −6.3		Nonlinear trend % diff = −3.4	

	Percentage Believing Courts Too Harsh, OK (COURTS)					
a. Percentage diff = 0	>.05	Accept	>.05	Accept	>.05	Accept
b. Percentage diff = constant	—	—	—	—	—	—
c. Difference, *a* − *b*	—	—	—	—	—	—
d. Percentage diff = linear trend	—	—	—	—	—	—
e. Difference, *b* − *d*	—	—	—	—	—	—
Final	No difference % diff = 0		No difference % diff = 0		No difference % diff = 0	
	Percentage Opposing Capital Punishment (CAPPUN)					
a. Percentage diff = 0	>.05	Accept	>.05	Accept	>.05	Accept
b. Percentage diff = constant	—	—	—	—	—	—
c. Difference, *a* − *b*	—	—	—	—	—	—
d. Percentage diff = linear trend	—	—	—	—	—	—
e. Difference, *b* − *d*	—	—	—	—	—	—
Final	No difference % diff = 0		No difference % diff = 0		No difference % diff = 0	

NOTES: All results based on χ^2 tests with appropriate degrees of freedom (see Taylor 1980); n.s. = not significant. Periods covered by the series are 1959–85 for GUNLAW, 1965–85 for COURTS, and 1960–85 for CAPPUN.

INTERCOHORT COMPARISONS

When the adjusted percentage differences are plotted over time and evaluated statistically (table 4), the gun control question shows no linear trends toward either increasing or decreasing cohort differences and an average adjusted percentage difference of between −9.1 and −3.4. This means that the oldest cohort gives slightly *more* liberal responses than the other three cohorts, but there is no evidence of either increasing or declining differences over time. In the cases of the youngest/oldest and middle-young/oldest comparisons, these differences are significant; for the third comparison, differences in gun law attitudes seem to follow a curvilinear pattern.

For the court and capital punishment questions, none of the adjusted percentage comparisons shows any significant difference. In each of the six comparisons, the fluctuations in differences are too random and not significantly different from zero to support even a constant difference model. All nine comparisons between the oldest and younger cohorts show the oldest cohort to be changing its attitudes on gun control, the courts, and capital punishment in much the same way as each of the younger cohorts.

The graphs of cohort differences in liberal sentiment show the kind of fluctuation over time observed in the standardized cohort trendlines. As seen in figure 3 (top), comparisons regarding gun control attitudes change over time, but mostly around an axis below zero. This is consistent with the negative average differences found in the first row of table 4. Attitudes about the courts and capital punishment (fig. 3) show similar fluctuations over time, but around zero, again reflecting the statistical analyses reported in table 4.[7]

Discussion

The findings reported in this paper are important for two reasons. The first concerns issues specifically related to aging, and the second relates more generally to attitudes about law and order. We consider each briefly in turn.

---

[7] The three comparisons depicted in figure 3*b* might at first appear to show a general negative slope, revealing a change from positive differences in the early years to negative differences during the later years. The statistical analyses, following Taylor 1980, weighted the early comparisons less because of the quota sampling used in all or part of each survey before 1977, and this might explain the middle-young/oldest results. The youngest/oldest and middle-oldest/oldest comparisons support a no difference model, however, even when the early surveys are weighted the same as the later ones.

AGING AND ATTITUDE CHANGE

The primary objective of this study was to determine if support existed for the notion that aging is associated with a tendency toward the adoption of increasingly conservative attitudes. In an attitude domain marked by a conservative shift in public opinion, the aging-conservatism hypothesis would predict that older cohorts lead the conservative wave, showing evidence of earlier change and producing widening cohort differences. While the data did confirm that overall changes in attitudes about law and order have been in a conservative direction, conservative shifts were generally in evidence for each of the four cohorts examined in the analysis.

An even more compelling refutation of the aging-conservatism argument comes from the results of the analysis of trends in cohort differences. Comparing the oldest cohort with each of the others on the three law and order items shows that either constant difference or zero difference models best fit these trends. In other words, the conservative trend is no more pronounced for the oldest cohort than for the other cohorts. From this evidence, we therefore conclude that the propensity to adopt conservative attitudes does not vary across cohorts. Rather, the observed patterns suggest that each of the cohorts is being equally influenced by period effects emanating from changes occurring in the society at large. We will return to this point in the final section to consider the sources of these effects.

These findings are relevant to another notion about the relationship between aging and attitude change, one that focuses less on the direction of change than on its very existence. Along with the idea that aging is accompanied by increasingly conservative attitudes, a related supposition is that growing old tends to result in attitudinal rigidity. Regarding the "educability" of older persons, Freud, for example, seemed to be expressing this view when he noted that

> The age of patients has this importance in determining their fitness for psycho-analytic treatment; that, on the one hand, more or less above the fifties, the elasticity of mental processes, on which the treatment depends, is as a rule lacking—older people are no longer educable—and, on the other hand, the mass of material to be dealt with would prolong the treatment indefinitely. (cited in Horton 1982, 1)

Similarly, data from "The Myth and Reality of Aging in America" study (National Council on the Aging 1975) suggest the prevalence of this notion among the general public. Among persons 18–64 years of age, only 19% believed that most people over 65 were very open-minded and adaptable, and only 34% of persons 65 and over thought that most persons over 65 were very open-minded and adaptable. In

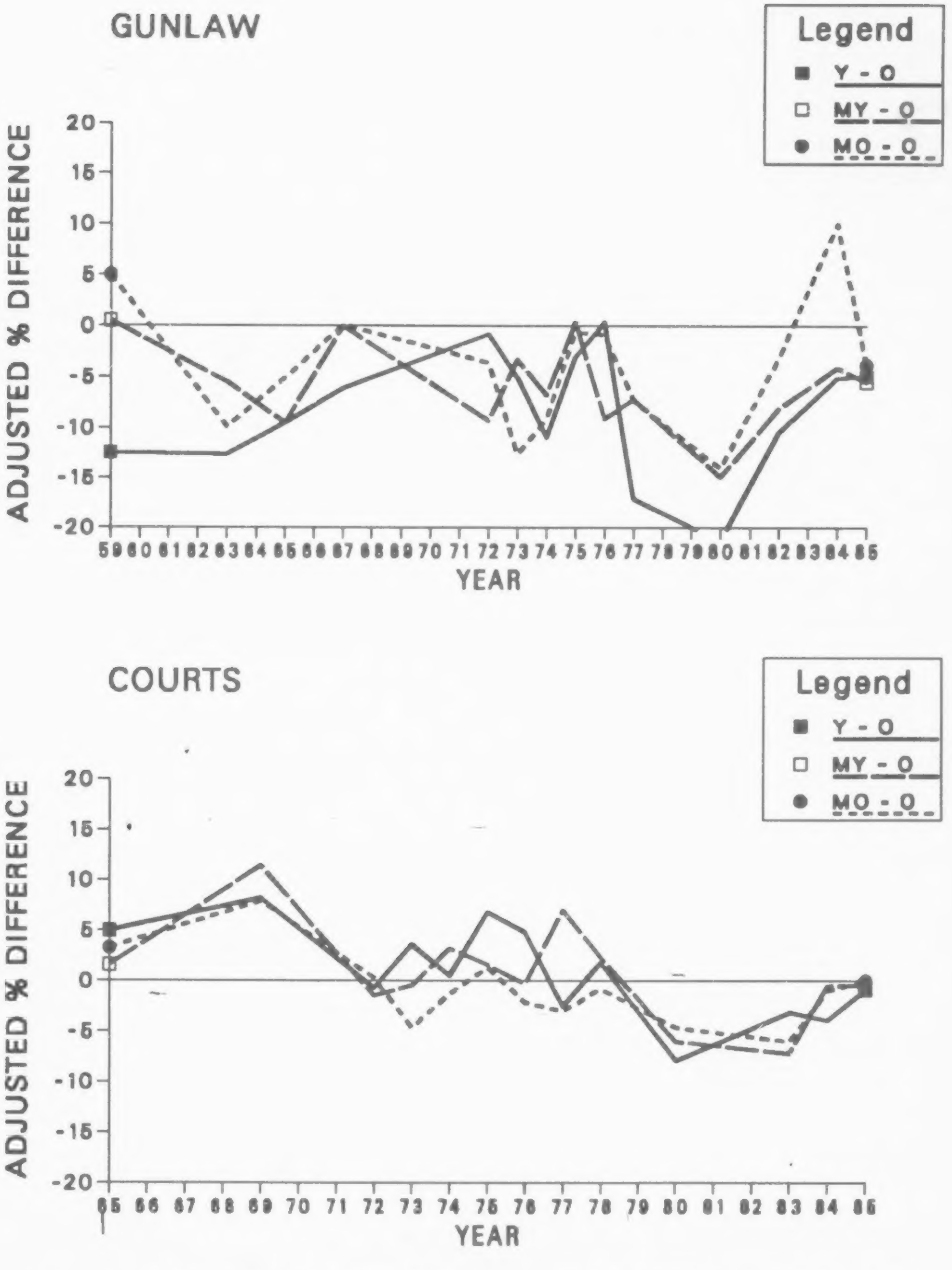


Figure 3. Trends in Cohort Differences in Law and Order Attitudes, Adjusted for Sex and Education Composition Differences (*top:* Gun Laws, *bottom:* Courts)

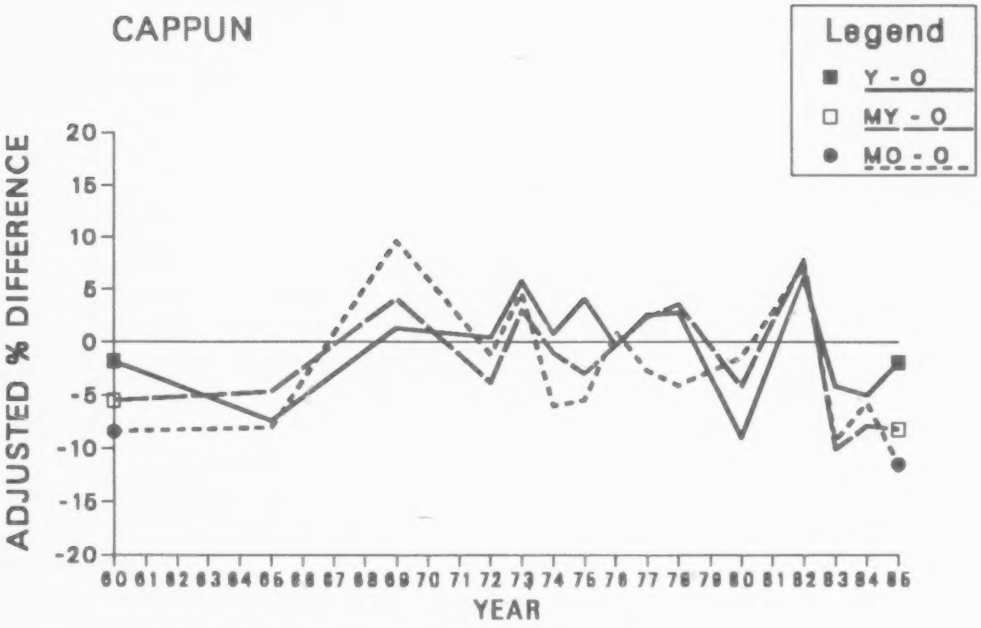


Figure 3 (*continued*). Trends in Law and Order Attitudes by Cohort and Year, Adjusted for Sex and Education Composition Differences (Capital Punishment)

its crudest form, this view is perhaps conveyed most clearly in the familiar saying "you can't teach old dogs new tricks."

When the findings from the present study are considered along with the results of other studies of aging and attitude change (e.g., Cutler and Kaufman 1975; Danigelis and Cutler 1989; Smith 1981; Thornton, Alwin, and Camburn 1983), the evidence about the presumed connection between aging and attitudinal rigidity seems clear and convincing. These investigations consistently demonstrate that, when change in public opinion is occurring in the society at large, older cohorts participate in the change, the direction in which their attitudes change is the same, and the rate of change (as we have found here) is more often than not similar across cohorts. In instances where little or no change is found among older cohorts (e.g., Cutler 1985), a lack of change in attitudes over time also characterizes the younger cohorts.

Thus, aging does not appear to be associated with an invariable tendency toward increasingly conservative attitudes, nor does it inevitably result in attitudinal rigidification. Aging may bring with it a variety of physiological, psychological, and social changes, but the long-held myth that aging is somehow "naturally" associated with an increasingly conservative set of ideas is clearly contradicted in the realm of law and order issues. While we have found elders to have become more conservative, so too have younger cohorts. The course of attitude change with aging, then, seems best predicted by the wider trends in public opinion occurring at that time.

THE SOCIAL SALIENCE OF CRIME

In their ambitious work on crime and punishment in the United States, Stinchcombe et al. (1980) emphasize "the salience of crime" as a key mediating variable between a variety of background predictors and attitudes toward the kinds of law and order questions we have been discussing. Key indicators of the salience of crime are "[f]ear, victimization, the lack of neighborhood safety, and seeing crime as a national problem" (Stinchcombe et al. 1980, 144). Given the nature of the law and order questions that we have examined and the consistency of findings across cohorts, social, rather than personal, salience would seem to be a more useful concept to explain our results. Alone among the salience indicators Stinchcombe and his colleagues used, seeing crime as a national problem, or "public interest in crime" (Stinchcombe et al. 1980, 24), measures *social*, as opposed to *personal*, salience.

The actual operational definition of this public interest is harder to find, according to Taylor, Scheppele, and Stinchcombe (1979, 423):

> Salience is only partly a result of changes in objective phenomena (e.g., crime rates). It is clearly influenced by media, political campaigns and all of the other diffuse sources of cultural information and cultural change which make detailed statements of certainty in public opinion research so hard to come by.

While Caplow's (1982) comparison of General Social Survey and Middletown data supports the idea that national media are important producers of change in public opinion, Combs and Comer (1982) find that television and newspapers are possibly important in accounting for the capital punishment attitudes of blacks, but not of whites. Furthermore, Stinchcombe et al.'s analysis of trend data leads them to conclude that "the great surge of public interest [salience] corresponds to the great surge in crime much better than to the periodic surges of media attention to crime and criminals" (1980, 24–25; also see Rankin 1979).

Acknowledging limitations of data and contradictory beliefs about the relevance of different aspects of the social salience of crime, we nevertheless can draw two conclusions from our findings. First, the trend toward more conservative attitudes regarding law and order among Americans involves young and old cohorts to a similar degree. While their starting points may differ, depending on which indicator of law and order attitudes is used, all four cohorts studied generally became more conservative, and at the same rate, thereby refuting the aging-conservatism argument. Second, to the extent that this study covers a particular period of time involving specific events of national importance (e.g., the Supreme Court decision on capital punishment, a number of assassinations, the rising crime rate), the results suggest a period, as opposed to a cohort or aging, effect in operation. Whether it is the crime rate, the media, or some combination of objective and subjective indicators of social salience that has produced the results is a question that is far from resolved, but the evidence for the existence of a national period effect is persuasive.

References

Alston, Jon P. 1976. "Japanese and American Attitudes toward the Abolition of Capital Punishment." *Criminology* 14:271–76.

Alston, Letitia T. 1986. *Crime and Older Americans*. Springfield, IL: Charles C. Thomas.

Berelson, Bernard, and Gary A. Steiner. 1964. *Human Behavior: An Inventory of Scientific Findings*. New York: Harcourt, Brace and World.

Campbell, Angus. 1971. "Politics through the Life Cycle." *Gerontologist* 11:112–17.

Campbell, John C., and John Strate. 1981. "Are Old People Conservative?" *Gerontologist* 21:580–91.

Caplow, Theodore. 1982. "Decades of Public Opinion: Comparing NORC and Middletown Data." *Public Opinion* 5:25–36.

Combs, Michael W., and John C. Comer. 1982. "Race and Capital Punishment: A Longitudinal Analysis." *Phylon* 43:350–59.

Coale, Ansley. 1964. "How a Population Ages or Grows Younger." In *Population: The Vital Revolution*, ed. Ronald Freedman, 47–58. Garden City, NY: Anchor Books.

Commission on Population Growth and the American Future. 1972. *Population and the American Future*. New York: New American Library.

Cutler, Neal E. 1977. "Demographic, Social-Psychological, and Political Factors in the Politics of Aging: A Foundation for Research in 'Political Gerontology.' " *American Political Science Review* 71:1011–25.

Cutler, Stephen J. 1985. "Aging and Attitudes about Sexual Morality." *Ageing and Society* 5:161–73.

Cutler, Stephen J., and Robert L. Kaufman. 1975. "Cohort Changes in Political Attitudes: Tolerance of Ideological Nonconformity." *Public Opinion Quarterly* 39:69–81.

Danigelis, Nicholas L., and Stephen J. Cutler. 1989. "Old Dogs and New Tricks: Cohort Changes in Racial Attitudes." Department of Sociology, University of Vermont. Typescript.

Davis, James A. 1975. "Analyzing Contingency Tables with Linear Flow Graphs: D Systems." In *Sociological Methodology 1976*, ed. David Heise, 111–45. San Francisco: Jossey-Bass.

DeFronzo, James. 1979. "In Search of the Behavioral and Attitudinal Consequences of Victimization." *Sociological Symposium* 25:23–39.

Exter, Thomas, and Frederick Barber. 1986. "The Age of Conservatism." *American Demographics* 8:30–37.

Firebaugh, Glenn, and Kenneth E. Davis. 1988. "Trends in Antiblack Prejudice, 1972–1984: Region and Cohort Effects." *American Journal of Sociology* 94:251–72.

Glenn, Norval D. 1974. "Aging and Conservatism." *Annals of the American Academy of Political and Social Science* 415:176–86.

Glenn, Norval D. 1975. "Trend Studies with Available Survey Data: Opportunities and Pitfalls." In *Survey Data for Trend Analysis: An Index to Repeated Questions in U.S. National Surveys Held by the Roper Public Opinion Research Center*, ed. Social Science Research Council, 6–50. Williamstown: Roper Center.

Glenn, Norval D. 1980. "Values, Attitudes, and Beliefs." In *Constancy and Change in Human Development*, ed. Orville G. Brim, Jr., and Jerome Kagan, 596–640. Cambridge, MA: Harvard University Press.

Glenn, Norval D., and W. Parker Frisbie. 1977. "Trend Studies with Survey Sample and Census Data." In *Annual Review of Sociology*, ed. Alex Inkeles, James Coleman, and Neil Smelser, 79–103. Palo Alto: Annual Reviews.

Herzog, Anna Regula. 1979. "Attitude Change in Older Age: An Experimental Study." *Journal of Gerontology* 34:697–703.

Horton, Arthur M., Jr. 1982. *Mental Health Interventions for the Aging*. New York: Praeger.

Hyman, Herbert H. 1972. *Secondary Analysis of Sample Surveys: Principles, Procedures, and Potentialities*. New York: Wiley.

Lawton, M. Powell, Lucille Nahemow, Silvia Yaffe, and Steven Feldman. 1976. "Psychological Aspects of Crime and Fear of Crime." In *Crime and the Elderly: Challenge and Response*, ed. Jack Goldsmith and Sharon S. Goldsmith, 21–29. Lexington, MA: D. C. Heath.

Mueller, John H., Karl F. Schuessler, and Herbert L. Costner. 1977. *Statistical Reasoning in Sociology*. 3d ed. Boston: Houghton Mifflin.

National Council on the Aging. 1975. *The Myth and Reality of Aging in America*. Washington, DC: National Council on the Aging.

Rankin, Joseph H. 1979. "Changing Attitudes toward Capital Punishment." *Social Forces* 58:194–211.

Riley, Matilda White. 1973. "Aging and Cohort Succession: Interpretations and Misinterpretations." *Public Opinion Quarterly* 37:35–49.

Ruggiero, Josephine, and Louise C. Weston. 1977. "Comparative Analysis of Male and Female Attitudes toward Tolerance of Nonconformists." *International Journal of Sociology of the Family* 7:77–85.

Schaie, K. Warner. 1967. "Age Changes and Age Differences." *Gerontologist* 7:128–32.

Schaie, K. Warner. 1983. *Longitudinal Studies of Adult Psychological Development*. New York: Guilford.

Schuman, Howard, and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. New York: Academic Press.

Sears, David O. 1981. "Life-Stage Effects on Attitude Change, Especially among the Elderly." In *Aging: Social Change*, ed. Sara B. Kiesler, James N. Morgan, and Valerie Kincade Oppenheimer, 183–204. New York: Academic Press.

Sengstock, Mary C., and Jersey Liang. 1979. "Attitudes of the Elderly toward the American Legal System." Paper presented at the 1979 Gerontological Society of America, Philadelphia.

Smith, A. Wade. 1981. "Racial Tolerance as a Function of Group Position." *American Sociological Review* 46:558–73.

Smith, Tom W. 1985. "Atop a Liberal Plateau? A Summary of Trends Since World War II." *Research in Urban Policy* 1:245–57.

Stinchcombe, Arthur L., Rebecca Adams, Carol A. Heimer, Kim Lane Scheppele, Tom W. Smith, and D. Garth Taylor. 1980. *Crime and Punishment: Changing Attitudes in America*. San Francisco: Jossey-Bass.

Stouffer, Samuel A. 1955. *Communism, Conformity, and Civil Liberties*. New York: Wiley.

Taylor, D. Garth. 1980. "Procedures for Evaluating Trends in Public Opinion." *Public Opinion Quarterly* 44:86–100.

Taylor, D. Garth, Kim Lane Scheppele, and Arthur L. Stinchcombe. 1979. "Salience of Crime and Support for Harsher Criminal Sanctions." *Social Problems* 26:413–24.

Thomas, Charles W., and Samuel C. Foster. 1976. "A Sociological Perspective on Public Support for Capital Punishment." In *Capital Punishment in the United States*, ed. Hugo Adam Bedau and Chester M. Pierce, 172–89. New York: AMS Press.

Thomas, Charles W., and Robert G. Howard. 1978. "Public Attitudes toward Capital Punishment." In *Justice and Corrections*, ed. Norman Johnston and Leonard D. Savitz, 407–17. New York: Wiley.

Thornton, Arland, Duane F. Alwin, and Donald Camburn. 1983. "Causes and Consequences of Sex-Role Attitudes and Attitude Change." *American Sociological Review* 48:211–227.

Vidmar, Neil, and Phoebe C. Ellsworth. 1976. "Public Opinion on the Death Penalty." In *Capital Punishment in the United States*, ed. Hugo Adam Bedau and Chester M. Pierce, 125–51. New York: AMS Press.

Warr, Mark, and Mark Stafford. 1984. "Public Goals of Punishment and Support for the Death Penalty." *Journal of Research in Crime and Delinquency* 21:95–111.

Williams, J. Sherwood, Charles W. Thomas, and B. K. Singh. 1983. "Situational Use of Police Force: Public Reactions." *American Journal of Police* 3:37–50.

Yin, Peter. 1985. *Victimization and the Aged*. Springfield, IL: Charles C. Thomas.

ESTIMATING COMMUNITY STANDARDS: THE USE OF SOCIAL SCIENCE EVIDENCE IN AN OBSCENITY PROSECUTION

DANIEL LINZ,
EDWARD DONNERSTEIN,
KENNETH C. LAND,
PATRICIA L. McCALL,
JOSEPH SCOTT,
BRADLEY J. SHAFER,
LEE J. KLEIN, and
LARRY LANCE

Abstract Elements of the legal test for obscenity of sexually explicit material indicted in a criminal case are examined. A cross-section of residents of Mecklenburg County (Charlotte, NC) were randomly assigned to view either one of the sexually explicit films and the sexually explicit magazine charged in the criminal case, or a control film. Before and after the viewing, residents judged the materials' appeal to a prurient interest (a shameful, morbid, unhealthy interest in sex) and patent offensiveness (community tolerance for such material). The results indicated that the respondents felt that the films and magazine *did not* appeal to a shameful, morbid, or unhealthy interest in sex, nor did they perceive these materials as going beyond the level of tolerance regarding depictions of sexual conduct for the average adult in that community. A lower percentage of subjects

DANIEL LINZ is Associate Professor of Communication at the University of California, Santa Barbara. EDWARD DONNERSTEIN is Professor of Communication at the University of California, Santa Barbara. KENNETH C. LAND is Professor of Sociology at Duke University. PATRICIA L. MCCALL is a graduate student in the Department of Sociology at Duke University. JOSEPH SCOTT is Associate Professor of Sociology at Ohio State University. BRADLEY J. SHAFER and LEE J. KLEIN are attorneys at law, Klein & Shafer P. C., Okemos, MI. LARRY LANCE is Associate Professor of Sociology at the University of North Carolina, Charlotte. The authors would like to thank Steven Penrod, Neil Malamuth, Donald Rubin, and William Thompson for their helpful comments on earlier drafts of this manuscript.

Public Opinion Quarterly 55:80–112 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-01/$2.50

thought the community tolerated the materials they had just viewed than when they were asked to report on what they *personally* tolerated. Fewer people felt the films appealed to a shameful, morbid, or unhealthy interest in sex *after* they had an opportunity to see them than before viewing. The advantages of providing jurors in obscenity cases with information about community standards based on summations of personal tolerance for materials actually charged in these cases, rather than hypothetical judgments about the community and obscenity, is discussed.

Introduction

There are few in-depth studies of the public's tolerance of pornographic or "obscene" materials. Existing studies tend to be from national surveys that only broadly describe respondents' "concern" about the pervasiveness of pornography in U.S. society (e.g., Yankelovich, Clancy, and Schulman 1986; Donnerstein and Bates 1987). Studies of trends in public opinion on permissiveness toward pornography suggest a complex pattern of shifting levels of tolerance. Some investigators (e.g., Smith 1987) have concluded that support for government regulation of pornography rose rapidly from 1975 to 1977 and then fell between 1977 and 1982. Comparisons of trends from the mid-1970s to the mid-1980s do suggest a slight decrease in public tolerance. These national studies, while interesting, are of limited value from a legal perspective because of the special requirements of obscenity law in the United States.

While the First Amendment of the U.S. Constitution protects all communicative material, including books, magazines, and films, and while these protections further extend to material containing nudity and sexually frank conduct, these constitutional protections do not extend to that which is legally "obscene." Thus, the sale and dissemination of obscene material may be criminally proscribed by law.[1] However, as described by the U.S. Supreme Court, obscenity and constitutionally protected material are separated by only a "dim and uncertain line" (*Bantam Books* 1963).

The decision of whether material is or is not "obscene" is a task that may be left to either a single judge or, more often, to a jury (the judge or jury is thus the "trier of fact" as to the issue of obscenity). Both must apply the proper legal standards in arriving at a just conclusion. While individual states, in theory, are free to adopt their own standards for the determination of obscenity, they cannot subject to criminal proscription that material which is protected, in the interpretation of the Supreme Court, by the First Amendment of the U.S. Constitution. The current legal standard of that Court to be applied in deciding whether material may be considered to be "obscene" is the tripartite test announced in the case of *Miller v. California* (1973), more commonly referred to as simply the "Miller Test." This test, in general, is statutorily codified by most states to be their criminal "obscenity" statute.

---

[1] Mere possession in one's own home, with the exception of child pornography (*Osborne v. Ohio* 1969), cannot be subject to criminal prosecution due to the overriding privacy concerns (*Stanley v. Georgia* 1969).

Pursuant to the *Miller* decision, in order to find material to be "obscene," the trier of fact (be it a judge or a jury) must determine (*a*) whether the average person, applying contemporary community standards, would find that the work, taken as a whole, appeals to the prurient interest; (*b*) whether the work depicts or describes, in a patently offensive way, sexual conduct specifically defined by applicable state law;[2] and (*c*) whether the work, taken as a whole, lacks serious literary, artistic, political, or scientific value. While the three-part test announced in *Miller* only discussed "community standards" in relation to the first prong of the test ("prurient interest"), subsequent decisions by the Supreme Court have clarified that the second prong ("patent offensiveness") is, similarly, to be judged with regard to "contemporary community standards," while the remaining element ("serious value") is not, but is, instead, to be judged with reference to the "reasonable person" (*Smith v. United States* 1977; *Pope v. Illinois* 1987). Therefore, obscenity is to be judged not by the personal opinions of the jurors, but, rather, by the jurors utilizing the concepts of the "average person applying contemporary community standards" and the "reasonable person."

In *Miller*, the Supreme Court specifically rejected the concept of a nationwide "contemporary community standard," and, instead, left to the individual states the determination of the scope of the relevant "community" by which the questioned material was to be judged. Thus, states are free to adopt a statewide standard for the relevant contemporary community, a smaller geographical boundary (i.e., a countywide standard), or, in fact, are allowed to not define the community at all, thereby allowing the trier of fact to draw upon their knowledge of the community or vicinage from which they come to decide how the "average person, applying contemporary community standards" would evaluate the particular material (*Hamling* 1974).

---

[2] While the Supreme Court did not specifically limit the types of sexual conduct that could be found to be patently offensive, it did provide such examples as representations or descriptions of ultimate sexual acts, normal or perverted, actual or simulated, and representations and descriptions of masturbation, excretory functions, and lewd exhibition of the genitals (*Miller* 1973, 25).

The purpose of this study was to ascertain the contemporary community standards of adults in Mecklenburg County, North Carolina, regarding sexually explicit material, both generally and with particular regard to five videotape films and a magazine that were subject to criminal indictments as allegedly being "obscene," and to determine, given those standards, whether the materials either appealed to a prurient interest or were patently offensive. The concept of "prurient interest" was defined in this study, pursuant to Supreme Court decisions, as being an "unhealthy, shameful, or morbid interest in sex" (*Roth* 1957; *Brockett* 1985). "Patent offensiveness" is considered to be descriptions or representations that go "substantially beyond the customary limits of candor" in such matters (*Roth* 1957, 478, fn. 20). This has been generally interpreted by courts to be the level of "tolerance" or "acceptance" of such materials in the community. Given that the North Carolina standard jury instructions regarding obscenity utilize "tolerance," this study defined "patent offensiveness" in terms of the level of tolerance for such materials. This study, as described below, however, can just as easily be conducted utilizing an "acceptance" standard if local interpretation so requires. While the substantive findings of this study may be limited to the county in which the data were collected, the methods used for determining "contemporary community standards" and the "prurient appeal" and "patent offensiveness" of sexually explicit materials and the implications for the introduction of this form of scientific evidence by experts in obscenity trials have, we maintain, more general application.

EXPERT TESTIMONY IN OBSCENITY CASES

Under most circumstances, the prosecution in a criminal obscenity action is not constitutionally required to introduce affirmative expert testimony in order for the trier of fact to determine whether or not materials are obscene. This is so because the Supreme Court has held that the materials themselves are the "best evidence of what they represent," and that, in general, no such expert testimony is necessary in order for the prosecution to meet its burden of proof of the obscenity of charged materials (*Paris Adult Theatre* 1973).[3] Thus, it is assumed that the trier of fact will know, or can determine on its own (beyond a reasonable doubt), the relevant community standards.

The Supreme Court has, however, further noted that the defendant in a criminal obscenity prosecution should be free to introduce "appropriate" expert testimony at trial[4] (*Kaplan* 1973). In reaching this conclusion, the Court cited Justice Frankfurter's concurring opinion in the earlier decision of *Smith v. California* (1959) in which he noted:

---

[3] However, as noted by the Ninth Circuit U.S. Court of Appeals, while the materials may be the "best evidence of what they represent, . . . the materials will not supply any information as to the community standards by which they are to be judged" (*United States v. 2,200 Paperback Books* 1977, 566 and 570).

> The uncertainties pertaining to the scope of scienter requisite for an obscenity prosecution and the speculative proof that the issue is likely to entail, are considerations that reinforce the right of one charged with obscenity—a right implicit in the very nature of the legal concept of obscenity—to enlighten the judgment of the tribunal, be it the jury or as in this case the judge, regarding the prevailing literary and moral community standards and to do so through qualified experts . . . for community standards or the psychological or physiological consequences of questioned literature can as a matter of fact hardly be established except through experts. . . .
>
> . . . Since the law through its functionaries is "applying contemporary community standards" in determining what constitutes obscenity . . . it surely must be deemed rational, and therefore relevant to the issue of obscenity, to allow light to be shed on what those "contemporary community standards" are. Their interpretation ought not to depend solely on the necessarily limited, hit-or-miss, subjective view of what they are believed to be by the individual juror or judge. It bears repetition that the determination of obscenity is for juror or judge not on the basis of his personal upbringing or restricted reflection or particular experience of life, but on the basis of "contemporary community standards." (*Smith* 1959, Frankfurter, J., concurring, 164–66)

USING PUBLIC OPINION POLLS TO ASSIST THE JURY IN DETERMINING COMMUNITY STANDARDS

Expert testimony on community standards has often included submission of results from public opinion polls.[5] A recent case that fully discusses these matters is the decision of *Saliba v. State of Indiana* (1985). After reiterating that jurors are not instructed to evaluate obscenity based on their personal opinions, the court said:

> . . . expert testimony based on a public opinion poll is uniquely suited to a determination of community standards. Perhaps no other form of evidence is more helpful or concise: "A properly conducted public opinion survey adequately insures a good measure of trustworthiness, and its admission may be necessary in the sense that no other evidence would be as good as survey evidence or perhaps even obtainable as a practical matter." (*Saliba* 1985, 1185)

---

[4] The distinction between allowing expert witness testimony on behalf of the defendant in a criminal obscenity prosecution, while not requiring such evidence by the prosecution for its burden of proof, is based, in part, on the protections afforded by the Sixth and Fourteenth Amendments to the U.S. Constitution, which require that a defendant in a criminal action be afforded the opportunity to present a defense.

[5] See, for example, Lamont 1973; Schauer 1976; Glassman 1978; Beckett and Bell 1979.

The admissibility of a public opinion poll depends primarily upon what the court in *Saliba* termed certain "circumstantial guarantees of trustworthiness," including whether generally accepted survey techniques were used in conducting the poll and correct statistical methods were utilized in evaluating the results of the poll.[6] The court laid out a specific set of criteria, generally accepted by most social scientists as necessary to ensure a poll's reliability and validity, so as to establish admissibility. Most important among these is that the "relevant universe" is examined, that a representative sample is drawn from this universe, and that the sample, questionnaire, and interviews be designed in accordance with generally accepted scientific standards.

Because our study was to evaluate contemporary community standards in North Carolina, we were concerned with the North Carolina Supreme Court's views about the admissibility of expert witness testimony and, in particular, public opinion survey research. Two decisions of that court, *State of North Carolina v. Anderson* (1988) and *State of North Carolina v. Mayes* (1988), had addressed this area and had generally concluded that survey evidence would be admissible in a criminal obscenity prosecution as long as the study was conducted pursuant to accepted methodology, and the opinions garnered therefrom were relevant to the issues before the court. In *Anderson,* however, the court had excluded certain testimony that it perceived was not particularly directed nor limited to the specific type of materials subject to prosecution, and because the particular study there had not focused on the materials alleged to be obscene, or on materials that were reasonably similar.

In addition to the concerns noted in *Anderson,* a Federal District Court had similarly issued an opinion excluding a public opinion study from evidence because the court there had found that the descriptive language of a telephone survey did not adequately convey "the impact of the visual image" and did not sufficiently apprise the interviewees "of the nature of the charged materials" (*United States v. Pryba* 1988, 1229–30). It was in particular regard to the concerns noted in the *Anderson* and *Pryba* decisions that the following study was designed.

---

[6] Admissibility of expert testimony and public opinion surveys is also subject to the fundamental prerequisite of legal relevancy and must also be admissible under the local rules of evidence. By these rules, expert testimony is properly admissible when it can assist the trier of fact to either understand the evidence or to determine the facts in issue (Federal Rule of Evidence, 702; adopted in most states). However, even *relevant* evidence can be excluded by the trial court if its probative value is substantially outweighed by the danger that it would be unfairly prejudicial or would have a tendency to either mislead the jury or to confuse the issues (Federal Rule of Evidence, 403).

The Present Study

The study reported here was proffered as evidence in the *State of North Carolina v. Cinema Blue of Charlotte, Inc.* (1989). In this study, an attempt was made to gather social science data to inform the jury as to whether the average adult in Mecklenburg County, NC, applying contemporary community standards, would find that five movies and one magazine alleged to be obscene by the state, either appealed to a prurient interest in sex or were patently offensive.

The project was informed by a concern for what types of information would be considered relevant, and thus legally admissible, according to the courts in *Saliba, Anderson*, and *Mayes*. Specifically, the project was designed to: (1) provide information gained from a scientifically selected sample of persons from the community of Mecklenburg County, (2) whose opinions about what the community tolerated, what they personally would tolerate, and what they personally found to appeal to a prurient interest were formed on the basis of (3) viewing the specific materials challenged in the case.

The study is strengthened by several design features. First, a sample of adults was randomly selected from the community. The use of such a probability procedure is the best way of assuring a representative sample of adults from the community. These adults were then randomly assigned to view either the allegedly obscene materials or control materials, and to then render opinions as to prurient appeal and patent offensiveness. As in any experiment, random assignment of subjects to test conditions and control conditions allows us to rule out the possibility that individual subject characteristics are responsible for the findings.

Second, the questions asked of participants in this study allow us to assess both the participant's view of what the "community" tolerates and whether or not he or she, personally, should be permitted to view these materials. It is possible that participants' perception of what is tolerated in the community differs substantially from their own reports of tolerance. This difference (should it occur) is, in itself, scientifically interesting. In addition, by aggregating the individual opinions across the sample—for both the study samples' guess about the community and their own views—we are able to estimate two forms of community tolerance. Both may provide jurors with an indication of the "contemporary community standards" that meaningfully augments their own judgments about what the community tolerates (although we will argue that the aggregation of personal judgments is the more useful of the two).

Third, the procedures employed in the present study permit us to compare subjects' opinions about community tolerance (i.e., patent offensiveness) and prurient appeal, both before and after viewing the materials charged in the case. Also, any shifts in opinion from pre- to postfilm viewing (should they occur) can be correlated with other variables measured in the study so that we might understand why subjects' opinions about the materials changed.

Finally, the participant recruitment procedures used here permit us to compare participants who volunteered for the study, knowing they may be asked to view sexually explicit materials, with those who declined to participate. With this information we are able to estimate the effects of any self-selection bias on responses to key questions.

FILM AND MAGAZINE VIEWING STUDY

A sample of Mecklenburg County residents were contacted by telephone. The sample was compiled through random digit dialing using a list of digits, not weighted by prefix, obtained from Survey Sampling Inc., Fairfield, Connecticut. Potential respondents who could not be reached were called back three times with next-day calling staggered by time. Once contacted, the respondents were administered an interview that involved four components: (1) selection of an adult respondent in the household 18 years old or over; (2) assessment of demographic characteristics including: age, education, race, and sex (the interviewers were instructed to recruit within quotas corresponding to the proportions within each of these demographic categories in the population [see n. 8]); (3) subject recruitment "to possibly view a movie . . . and fill out a questionnaire about it?" and (4) subject recruitment to possibly view "adult, X-rated sexually explicit films." Once respondents agreed to participate in the study, they were informed that they would receive $50.00 upon completion of the project, which, they were told, would last approximately two-and-a-half hours. Participants were then given the address of the study site.

A breakdown of response rates at each stage of subject recruitment is presented in table 1. A total of 614 adults completed the initial telephone interview. Of these, 380 respondents indicated that they would be willing to participate in a film-viewing project; 8 indicated they were "not sure," and 226 declined participation. From this group, a total of 284 respondents agreed to participate in the study once they were informed that they might be asked to watch an X-rated video. Of these respondents, 244 were scheduled for a film-viewing session. One hundred thirty-two subjects reported for a film-viewing session. One subject left once the video began, and two were dismissed with payment because all video viewing set-ups were being used simultaneously. A total of 129 subjects participated in the film-magazine evaluation study.

Table 1. Breakdown of Subject Loss at Each Stage of Film Evaluation Study

	N	Percentage Remaining		
Total numbers called	5,810	100		
Nonworking numbers	1,585	73		
No answer/no contact	867	58		
Business or government	613	47		
Ineligible county	8	47		
Eligible households	2,737		100	
Refusals	1,063		61	
Call-back at last contact	219		53	
Answering machine	199		46	
Busy signal at last contact	190		39	
Language or hearing problem	104		35	
Over quota for demographic category	320		23	
Miscellaneous	28		22	
Completed telephone interview	614			100
Refused to participate in a film-viewing project	226		14	63
Declined to view an X-rated video	104		10	46
Could not be scheduled for viewing session	40		9	40
Scheduled but cancelled by researchers	22		8	36
Failed to show for scheduled session	90		5	21
Miscellaneous	3		5	21
Total subjects viewing films	129			

We acknowledge that nonresponse in this study is higher than would be the case if respondents were simply asked their opinions over the telephone. However, the unique demands placed on respondents who must report for a laboratory study of this nature may preclude a lower nonresponse rate. In a subsequent section we assess the potential bias resulting from nonresponse at the most critical drop point—once respondents have been informed that they may be viewing X-rated materials.

ASSIGNING SUBJECTS TO FILM VIEWING CONDITIONS

The study began January 21, 1989, and was terminated on February 8, 1989. A seven-step procedure was followed with subjects: (1) consent form completion, (2) prefilm questionnaire administration, (3) movie viewing, (4) completion of postmovie questionnaire, (5) magazine viewing, (6) completion of magazine questionnaire, and (7) debriefing and payment.

Subjects reported to the viewing sessions not knowing if they would be asked to watch an X-rated film or a nonexplicit film. Subjects were randomly assigned to view either one of five sexually explicit films: *Unnatural Phenomenon 2* (*N* = 25), *The Big Switch* (*N* = 24), *How to Get A'head'* (*N* = 20), *Penetration 2* (*N* = 21), *Cumshot Revue Vol. 4* (*N* = 20), or a nonexplicit control film, *Nothing in Common* (*N* = 19).

OBTAINING INFORMED CONSENT

Prior to participating in the study, each subject assigned to an X-rated viewing condition was given a consent form to read over and sign. This document informed them that they would be asked to view a sexually explicit, X-rated adult video and a magazine of the same content.

PREFILM QUESTIONNAIRE ADMINISTRATION

Subjects were then given a questionnaire and instructed to take it to one of ten separate video viewing rooms and complete it. They were asked to indicate length of residence in Mecklenburg County, newspaper subscription, video movie consumption, cable subscription, political affiliation, political attitudes, religiosity, religious service attendance, attitudes about current events, education level, age, and religious affiliation. Most important, the following set of instructions were given and questions asked.

> The next few questions deal with X-rated videos and adult movies and magazines. The nudity and sex shown in these types of adult material include: nude bodies and close-up graphic depictions of a variety of sexual activities, such as sexual intercourse, ejaculation, bondage, oral sex, anal sex, group sex, and variations of these acts by adult performers.
>
> Do you think it *is* or *is not* tolerated in your community for the average adult to see adult movies, video cassettes, and magazines showing nudity and sex if they should want to?
>
> Tolerated 1 2 3 4 5 6 7 8 Not Tolerated
>
> Do you believe you *should* be able to see any such showing of actual sex acts in adult movies, video cassettes, or magazines if you should want to?
>
> Definitely No 1 2 3 4 5 6 7 8 Definitely Yes

> Some adult movies, video cassettes, and magazines show actual sex acts in great detail and with close-ups of sexual organs. Would viewing this type of material appeal to an unhealthy, shameful, or morbid interest in sex?
>
> Definitely Yes 1 2 3 4 5 6 7 8 Definitely No

FILM VIEWING

Upon completing the prefilm questionnaire, subjects were instructed to return it to one of the staff members. The subject was then taken to the viewing room and shown how to turn off the video cassette recorder (VCR). Subjects were instructed at that time to turn off the VCR once they had viewed the videocassette and report to the staff person. The staff person then started the VCR and left the room. All subjects viewed the film alone, and their viewing was timed to ensure that they had viewed the film in its entirety.

POSTFILM QUESTIONNAIRE

When subjects returned from their movie viewing room, they were given a postfilm questionnaire and instructed to return to their room, alone, to complete it. This questionnaire assessed subject attitudes on a number of current issues, including gun control, censorship, and religious tolerance. Most important, they were given the following instructions and questions.

> In the first questionnaire you filled out, we asked you about X-rated videos, adult movies, and adult magazines. We said that the nudity and sex shown in these adult materials include: nude bodies and close-up graphic depictions of a variety of sexual activities. This includes: sexual intercourse, ejaculation, bondage, oral sex, anal sex, group sex, and variations of these acts by adult performers.
>
> Was the movie you saw today what you expected to see, given this definition of X-rated videos, adult movies, and magazines?
>
> Definitely No 1 2 3 4 5 6 7 8 Definitely Yes
>
> The next few questions deal *only* with the X-rated video that you saw today. The scenes you saw in this film may have included nude bodies and close-up graphic depictions of a variety of sexual activities, including: sexual intercourse, ejaculation, oral sex, anal sex, group sex, and variations of these acts by adult performers.
>
> Have you ever viewed a film or video like the one you just saw?
>
> Yes ____ No ____

> If you answered *Yes*, about how many times in the last five years? _____

This question was followed by questions on community tolerance and personal tolerance that were similar to those in the pretest. The appeal to prurient interest question was not preceded by the sentence concerning adult movies and videotapes in general, and reference was made, rather, to the particular film just viewed.

MAGAZINE VIEWING AND MAGAZINE QUESTIONNAIRE ADMINISTRATION

After completion of the postfilm questionnaire, subjects were given a packet that contained the magazine *Pregnant and Sexy* and a questionnaire. They were instructed to take the packet back to their room, look at the magazine, and complete the questionnaire. This questionnaire included questions similar to the postfilm questionnaire, except that the questions on film viewing were modified to apply to magazine viewing, subjects were not asked about their definition of X-rated videos, and subjects were given the following instructions.

> The next few questions deal with the X-rated magazine that you just viewed. You may feel exactly the same way about the magazine you just examined as you did about the video that you viewed. If you feel the same way, it is okay to give the same answers you gave on the questionnaire about the video. If you feel differently about the magazine, you can, of course, indicate that also.

SUBJECT PAYMENT AND DEBRIEFING

Once the postmagazine questionnaire had been completed, subjects were paid $50.00 and thanked. Subjects were then asked if they had any questions about the study. The research assistants were instructed to answer each question honestly and completely.

CONTROL FILM SUBJECTS

Control subjects did not participate in the magazine-viewing phase of the study, and they viewed a nonexplicit film (*Nothing in Common*). Their postfilm questionnaire did not include questions on whether the film they viewed met their definition of an "X-rated" film, and they were asked the prurient appeal question with the preface sentence in both the pre- and postfilm questionnaire. Otherwise, control subjects were treated in an identical manner to the sexually explicit film-viewing subjects.

Results

SUBJECT CHARACTERISTICS

On the average, subjects in the film and magazine evaluation sample have lived in Mecklenburg County for about 19 years, and the majority of people subscribe both to a daily newspaper and cable television. Over 50% of the persons in the sample consider themselves to be politically conservative to very conservative and over 80% say they are somewhat to very religious. Most had completed high school, over half had some college or completed a college education, and over half were married. The ages of subjects ranged from 18 to 69 years. The majority had an annual income over $30,000 per year, although at least 20% indicated a yearly income under $20,000 per year.[7]

INITIAL ASSESSMENT OF SUBJECT SAMPLE SELECTION BIAS

Despite our efforts to establish quotas within demographic categories, some respondents were more willing to participate in the film-viewing study than others once they had been informed over the telephone that they may be asked to watch an X-rated film. Table 2 shows the number and percentage of respondents, broken down by the primary demographic variables—age, sex, race, and education—who agreed and who did not agree to participate after they were told they may be asked to view X-rated materials. By comparing the percentages in the "agree to participate" column for each demographic category, we see that nonwhites were more likely to volunteer for the study than whites ($\chi^2 = 14.25$, $df = 1$, $p < .05$), males more than females ($\chi^2 = 10.36$, $df = 1$, $p < .05$), and younger persons more likely to volunteer than older persons ($\chi^2 = 8.36$, $df = 1$, $p < .05$). Respondents with a moderate level of education were slightly more likely to volunteer than respondents with either higher or lower levels of education, although differences between categories were nonsignificant ($\chi^2 = 3.52$, $df = 2$, $p =$ n.s.).

As a further check of response bias, we compared the percentage

---

[7] Charges were dropped against three of the films prior to trial. The films *Cumshot Revue Vol. 4* and *The Big Switch* were, however, prosecuted. A total of 44 subjects viewed these films. On the average, subjects in that sample lived in Mecklenburg County for about 18 years and the majority of people subscribe both to a daily newspaper and cable television. Forty-five percent of the persons in the sample consider themselves to be politically conservative to very conservative and over 80% indicate they are somewhat to very religious. Most had completed high school, over half had some college or completed a college education, and over half were married. The ages of subjects ranged from 18 to 62 years. The majority had an annual income over $30,000 per year, although at least 20% indicated a yearly income under $20,000.

Table 2. Comparison of Subjects Who Agreed and Did Not Agree to Participate in the Study

	Participate	Not Participate	Total
Race			
White	70% (186)	30% (81)	100% (267)
Black	88 (97)	12 (13)	100 (110)
Sex			
Male	85 (115)	15 (21)	100 (136)
Female	70 (168)	30 (73)	100 (241)
Education			
12 years or less	73 (101)	27 (37)	100 (138)
13–15 years	82 (80)	18 (17)	100 (97)
16 or more	72 (102)	28 (39)	100 (141)
Age			
18–39	80 (177)	20 (43)	100 (220)
40 or older	67 (105)	33 (51)	100 (156)

NOTE: *N*s of subgroups are in parentheses.

of subjects who volunteered for the present study with participants in a telephone survey of community standards conducted in 1987 that employed a substantially larger sample of Mecklenburg County residents.[8] Table 3 presents these comparisons by the four demographic

[8] A random digit dialing telephone sample (Survey Sampling, Inc., Fairfield, Connecticut) of Mecklenburg County residents was conducted in July, 1987, as part of a study of community standards in an unrelated legal proceeding. A total of 875 households were contacted. Two hundred six numbers resulted in incompletion due to: language difficulties (35), busy signal (38), answering machine (62), or "call backs" who were never reached (71). Sixteen respondents terminated the interview midstream, and 246 declined to be interviewed at all. A total of 407 respondents, 18 years old or older, completed the interview.

Respondents were questioned about length of residence in Mecklenburg County, television and newspaper consumption habits, political affiliation, religiosity, and a variety of other issues. In addition, level of education, income, race, sex, and age was assessed. Response formats for the questions used in the 1987 survey and the prefilm questions asked in this 1989 study were identical. Subjects were also given the same description of X-rated videos used in the film evaluation study and asked the three questions pertaining to tolerance in the community for adult movies, belief that the respondent should be able to see these materials, and appeal to an unhealthy, shameful, or morbid interest in sex.

Respondents were instructed to answer "Yes" or "No" for each question. Rather than answering "Yes" or "No," subjects in the film evaluation study were asked to circle a number between 1 and 8 anchored by the terms "Tolerated" and "Not Tolerated" for the first question and "Definitely No" and "Definitely Yes" for the second and third questions. In comparisons involving the 1987 survey and the 1989 film evaluation project data, we make the assumption that the response formats for the film evaluacategories. Statistical comparison of the cell entries supported the possibility, suggested above, of a volunteer bias in at least two of the demographic categories, sex ($\chi^2 = 3.22$, $df = 1$, $p < .10$), and age ($\chi^2 = 14.42$, $df = 2$, $p < .01$). Statistical comparisons of cell entries for race and education were nonsignificant ($\chi^2 = .17$, $df = 1$, $p =$ n.s., for race; $\chi^2 = .20$, $df = 2$, $p =$ n.s. for education).

As a final and more specific check on response bias, we compared our prefilm/magazine-viewing responses to the three key prefilm questions—(1) whether it is or is not tolerated in the community for adults to see adult magazines and movies, (2) whether or not the respondent should be allowed to see such materials, and (3) whether such materials appeal to a shameful, unhealthy, or morbid interest in sex—with responses to identical questions asked in the telephone survey conducted in 1987. This provides an index of whether our film/magazine-viewing sample was either more or less favorably disposed to viewing sexually explicit materials than the population as a whole before participating in the film-viewing phase of the study. Table 4 presents the percentages for responses to the three critical questions asked in the 1987 Survey with responses from the film study subjects before they viewed the X-rated materials. A comparison of the cells in a 2 × 2 chi-square analysis revealed statistically significant differences among the cells for question 1 ($\chi^2 = 9.08$, $df = 1$, $p < .01$), no differences for question 2 ($\chi^2 = 1.26$, $df = 1$, $p =$ n.s.), and significant differences among the cell entries for question 3 ($\chi^2 = 14.13$, $df = 1$, $p < .01$).

These comparisons indicate that subjects who volunteered for the film/magazine-viewing study could be considered somewhat more favorably disposed to sexually explicit materials than the community at large on some dimensions, but less so on others. Film/magazine-viewing study participants were more likely to indicate that sexually explicit materials were tolerated in the community than the respondents in the 1987 baseline telephone survey. Yet, respondents in the two samples shared similar levels of approval about the extent to which they personally should be permitted to view the material themselves. However, subjects who volunteered for the study were substantially *more* likely to say that materials that portray explicit sex appealed to

---

tion study can be collapsed into dichotomous variables and the responses from both studies can be reliably compared to one another. In doing so we recognize several differences between the formats that may make them somewhat incompatible. The "yes" or "no" response format used in the telephone survey may capture attitude position, but not the intensity with which the attitude is held. The scaled response format used in the film evaluation study, on the other hand, may capture both position and extremity (or may even confound the two dimensions; see Coverse and Presser 1986).

Table 3. Comparison of Four Demographic Characteristics for Film Study Participants and 1987 Survey Respondents

	Race		Sex		Age			Education		
	Black	White	Male	Female	18–29	30–49	50+	≥12	13–15	16+
Film study	24%	76%	55%	45%	28.7%	55.8%	14.7%	32.4%	33.3%	34.1%
N	31	98	71	58	37	72	19	42	43	44
1987 survey	22.2%	77.8%	45.9%	54.1%	33.9%	40.6%	25.5%	34.2%	33.7%	32.1%
N	90	315	187	220	133	159	100	139	137	130

Table 4. Comparison of 1987 Survey Results, "Before" and "After" Film Viewing (X-rated film viewing subjects only) on Beliefs About Community Tolerance, Viewing Materials, and Appeal to Shameful and Morbid Interest in Sex

	1987 Survey	Before Film	After Film
	"Do you think it *is* or *is not* tolerated in your community for the average adult to obtain and see adult movies, video cassettes, and magazines showing nudity and sex if they want to?"		
Tolerated[a]	42.7%	59.1%	52.7%
N	156	65	58
Not tolerated	57.3%	40.9%	47.3%
N	209	45	52
	"Do you believe you *should* be able to see any such showing of actual sex acts in adult movies, video cassettes, or magazines if you want to?"		
Definitely yes	69.0%	74.5%	75.19%
N	265	82	86
Definitely no	31.0%	25.5%	24.81%
N	119	28	24
	"Some adult movies, videocassettes, and magazines show actual sex acts in great detail and with close-ups of sexual organs. Would viewing this type of material appeal to an unhealthy, shameful, or morbid interest in sex?"[b]		
Definitely yes	20.5%	43.6%	16.4%
N	71	48	18
Definitely no	79.5%	56.4%	83.6%
N	275	62	92

[a] Respondents to the 1987 survey answered "Yes" or "No" to this question. Subjects in the film evaluation study were asked to circle a number between 1 and 8, anchored by the terms "Tolerated" and "Not Tolerated." These subjects were also instructed to circle a number between 1 and 8, anchored by "Definitely No" and "Definitely Yes," for the second and third questions. In some cases, respondents to the 1987 survey volunteered the response "don't know." These rates were 10.3% for question 1, 5.7% for question 2, and 15.0% for question 3. Respondents indicating "don't know" are not considered here.

[b] The postviewing wording of this question was: "Did your viewing of this X-rated video appeal to any unhealthy, shameful, or morbid interest in sex that you might have?"

an unhealthy, shameful, or morbid interest in sex than were baseline respondents.

At least two statistical techniques are available that permit us to take this sampling bias, and any bias resulting from underrepresentation in a demographic category, into account for each of the major findings. We will discuss the results of the application of one of these techniques to our data following our report of the major findings.

SUBJECT RESPONSES AFTER FILM VIEWING AND COMPARISONS PRE- AND POSTFILM VIEWING

Column 3 of table 4 shows the responses of subjects to the three critical questions after viewing the films. (It is important to remember that subjects were instructed to answer these questions *only* with reference to the film they had viewed that day.) For these analyses, responses were summed across all of the X-rated film-viewing conditions. As can be seen from the table, a majority of subjects believe that it is tolerated in their community for the average adult to obtain and see the adult video cassettes charged in the case. Seventy-five percent believe they personally should be able to see any such showing of actual sex acts in adult movies, video cassettes, or magazines, and 16% believe viewing the type of material charged in the case appealed to an unhealthy, shameful, or morbid interest in sex.

Changes pre- and postfilm viewing (again, summed across the five sexually explicit film conditions) were also examined for each of the three outcome variables by comparing the mean differences from pre- to postviewing. Comparisons of pre- and postfilm subject responses on the belief that these materials are tolerated in the community showed no significant change from pre- to postfilm viewing (t = 1.41, df = 108, $p > .15$). Likewise, there was no statistically significant change in subjects' opinions about whether they should be allowed to see such materials (t = .51, df = 109, $p > .50$). There was, however, a statistically significant change in subjects' perception that the materials appealed to a shameful, unhealthy, or morbid interest in sex before and after viewing. Subjects were less likely to find the materials appealed to such interests after viewing (t = 6.75, df = 108, $p < .0001$). Forty-four percent of the subjects believed that such material appealed to morbid, shameful, or unhealthy interest in sex before seeing a film. After they viewed a film, the percentage dropped to 16%.

Computation of Pearson r correlations designed to "explain," post hoc, the shift toward finding the materials less likely to appeal to prurient interests were also undertaken. A change score reflecting the shift for each subject was computed and this score was then correlated with the demographic and attitudinal variables measured for all subjects, as well as the items designed to assess whether subjects had seen a film like the one they had just viewed at any time during the last five years. Only one variable was significantly correlated with the change score—the variable assessing the number of films like these subjects had seen before ($r\ [N = 110] = -.186, p < .05$, two-tailed). The sign of the coefficient indicates that the more films of this sort subjects have seen, the less they changed in their evaluation of the films from pre- to postviewing.

ESTIMATING THE EFFECTS OF SUBJECT SAMPLE SELECTION BIAS ON RESPONSES TO THE PRIMARY OUTCOME VARIABLES

As noted earlier, some types of respondents were more willing to participate in the film-viewing study than others, once they had been informed that they might be asked to watch an X-rated film. We noted that, of the 388 telephone interview respondents who completed the initial interview and did not definitely decline to participate in the film-viewing project, some 104 (100 with complete questionnaires) refused when informed of the possibility they would see an X-rated film. It is now appropriate to ask whether or not the failure to include opinions from these 100 telephone interview respondents who had refused has grossly distorted the percentages responding positively (or negatively) to the three obscenity questions in table 4. Stated in more technical terms, we may ask: Have the sample self-selection processes governing the willingness of individuals to participate in the film-viewing study—whether by the explicit race, sex, education, and age categories (see tables 2 and 3) or implicitly by unmeasured variables (such as religiosity or political conservatism which were not ascertained in the telephone recruitment interviews) or by personal positions of individuals on the dependent variables themselves—so truncated the frequency distributions observed in the film-viewing study that the percentage estimates in table 4 are grossly biased and inaccurate?

We applied the mixture modeling approach developed by Rubin (1977) to ascertain the extent to which the "after film" response distributions in table 4 could be biased by sample self-selection. (See the appendix for further details on this as well as an alternative method for estimating response bias.)

Table 5 reports the centers and interval limits (upper and lower bounds) for the subjective 95% intervals we obtained from applying the mixture modeling procedure to the three response variables. The centers of the subjective intervals in table 5 can be compared with the corresponding "after film" percentages reported in table 4 to ascertain how much the mean percentage responses differ after applying the

Table 5. Centers and Lower and Upper Bounds for 95% Bayesian Intervals on Responses to the "Community Standards" Questions

		Interval Bounds								
		$\theta_1 = 0.1$			$\theta_1 = 0.2$			$\theta_1 = 0.4$		
Dependent Variable	Interval Center	$\theta_2 = 0.1$	$\theta_2 = 0.2$	$\theta_2 = 0.4$	$\theta_2 = 0.1$	$\theta_2 = 0.2$	$\theta_2 = 0.4$	$\theta_2 = 0.1$	$\theta_2 = 0.2$	$\theta_2 = 0.4$
Tolerate in community, percent yes	57	51, 63	51, 63	50, 64	46, 68	46, 68	46, 68	36, 78	36, 78	36, 78
Should be able to see, percent yes	76	69, 84	69, 84	69, 84	61, 92	61, 92	61, 92	46, 100	46, 100	46, 100
Appeal to unhealthy interest in sex, percent yes	16	13, 18	13, 19	13, 19	12, 20	12, 20	11, 20	9, 23	9, 23	9, 23

correction for sample self-selection. The 95% intervals are reported for three values (0.10, 0.20, and 0.40) for each θ_1 and θ_2, representing successively more extreme hypotheses about how the regression coefficients and expected values on the (unobserved) dependent variables of the nonparticipants differ from the film-viewing participants. For example, for the question pertaining to whether participants believe that they should be able to see such actual sex acts, the upper and lower limits of the 95% interval for $\theta_1 = 0.10$ and $\theta_2 = 0.20$ are 69% and 84%, respectively, in table 5. This θ_1 specification says that the regression coefficients of the *Y* on *X* (unobserved) regressions for the participants are hypothesized to lie within an interval of plus or minus two (since we seek 95% intervals) times 0.10 around the estimated regression coefficients for the film-viewing participants—that is, that the participants and nonparticipants differ only slightly in terms of the relationships between their demographic characteristics and their attitudes on the dependent variables. Similarly, this θ_2 specification says that the expected value on the dependent *Y* for nonparticipants whose average values on the independent *X* variables are the same as the average *X*-values of the film-viewing participants is hypothesized to fall within an interval of plus or minus two times 0.20 around the expected value of *Y* for the film-viewing participants with average values on the independent variables. These θ_1 and θ_2 values then are used to compute the limits of the interval reported in the table within which we would expect the participant sample percentage responding "yes" to fall 95% of the time if there had been no telephone respondents who refused to participate when informed they might see an X-rated film. It can be seen that the 95% intervals in table 5 display more sensitivity to the hypotheses about differences in the θ_1 (regression) parameters than in the θ_2 (expected value) parameter.

It can be seen from table 5 that the 95% interval centers for the three response variables do not differ substantially from the corresponding "after film" percentages reported for the film-viewing study participants (table 4). That is, after correcting for possible sample self-selection bias, substantive conclusions about the percentages reported in table 4 are not greatly altered. Similarly, it can be seen from table 5 that extremely more radical and progressively less plausible assumptions about how nonparticipants differ on these three questions from participants are required to obtain 95% intervals that contain percentages on a lower or upper bound that move "in the opposite direction" or beyond the 50% mark from those reported in table 4. Indeed, for the "appeal to prurient interest" question, even our most extreme specifications on θ_1 and θ_2 do not produce 95% intervals that go beyond the 50% level. In sum, we cannot plausibly reject the conclusions drawn from the percentages reported in table 4 because of sample self-selection.

FURTHER SUBSTANTIATION OF THE FINDINGS

In addition to the foregoing evidence that the study did not plausibly suffer from a subject self-selection threat before viewing the films, we undertook additional analyses to further increase our confidence in the results we obtained after subjects had viewed the films. It is necessary to rule out the possibility that answering the questions about tolerance/patent offensiveness and prurient appeal twice, once in the prefilm questionnaire and again in the postfilm questionnaire, might have caused subjects to think more carefully about these issues than they normally would, and lead them to change their opinion as a result of this reflection (a testing-effect bias). Specifically, it may be the case that the significant change from pre- to postfilm viewing on the questionnaire item about unhealthy, shameful, or morbid interest in sex that we observed was the result of subjects giving this question further consideration from the time of completing the prefilm questionnaire to the time they completed the postfilm questionnaire, and not the result of changing their opinions because they had seen the film. To test this possibility, we compared the responses of the neutral-film control group subjects before and after the film viewing on this questionnaire item. The results of this statistical test revealed a nonsignificant change between pre- and postfilm evaluations for the control group on the prurient appeal question ($t = 1.15$, $df = 18$, $p > .26$).

POSTMAGAZINE QUESTIONNAIRE RESPONSES

Only subjects who viewed one of the X-rated films were shown the magazine *Pregnant and Sexy* and completed the postmagazine questionnaire; the control group did not see this magazine. The results indicate that, overwhelmingly, subjects felt that they personally should be permitted to view sexual acts like those depicted in the magazine (subjects scoring 5 or above on the 1–8 scale anchored "Definitely No" and "Definitely Yes" = 74%), and that the content of the magazine did not appeal to a morbid, shameful, or unhealthy interest in sex (subjects scoring 5 or above on the 1–8 scale anchored "Definitely No" and "Definitely Yes" = 22%). Subjects were divided evenly on the question of whether it was tolerated in their community for the average adult to obtain and see adult magazines showing such nudity and sex if they should want to (subjects scoring 5 or above on the 1–8 scale anchored "Tolerated" and "Not Tolerated" = 50%).

Discussion

The results of this study demonstrate the following: (1) the particular films and magazine viewed by the participants in this study, and charged in the case *State of North Carolina v. Cinema Blue of Charlotte, Inc.*, do not, for the majority of adult residents in Mecklenburg County, North Carolina, appeal to a self-reported shameful, morbid, or unhealthy (prurient) interest in sex; (2) there is a substantial shift in these opinions before and after viewing the films—far fewer people felt the films appealed to a prurient interest in sex *after* they had an opportunity to see them than before; (3) the particular films and magazine viewed by the participants were not, according to the majority of subjects, patently offensive—that is, they did not go substantially beyond the customary limits of candor (tolerance) in Mecklenburg County, NC, in depictions of sex; (4) a much lower percentage of people think other members of the community tolerate the films and magazine they just viewed than is the case when people are asked to report on what they personally tolerate, or should be allowed to view; and (5) statistical analyses designed to determine if these results were affected by sample selection bias (the tendency for men, younger persons, and persons already more tolerant of sexually explicit materials to be more likely to volunteer for the study) showed no plausible support for such contaminating effects.

The procedures and analyses employed in the study met the criteria for the type of information most likely to be considered relevant and, thus, legally admissible according to the courts in *Saliba, Anderson,* and *Mayes,* and corrects for the perceived failings as noted by the *Pryba* and *Anderson* courts. This particularly included obtaining information from a scientifically selected sample of persons from the community whose opinions were formed on the basis of viewing the *specific* materials challenged in the case.

More generally, the research technique described here has implications for the role of the social scientist in assisting jurors and judges in determining contemporary community standards toward patent offensiveness and prurient appeal of sexually explicit materials in this and other jurisdictions.

As noted in the introduction, the test for obscenity outlined in *Miller v. California* (1973) and adopted by nearly every state has several elements. More fully stated, the jury or judge must decide whether (1) the average person, (2) applying contemporary community standards would find that (3) the work, taken as a whole (4) appeals to the prurient interest; (5) whether the average person applying contemporary community standards would find that the work depicts or describes in a patently offensive way (6) sexual conduct specifically defined by state law; and (7) whether the work . . . lacks serious literary, artistic, and political or scientific value (Brigman 1978). We will consider elements 1 through 5 in light of the evidence presented in this study.[9]

DEFINING THE AVERAGE PERSON

Obscenity law often appears to presume that a properly constituted jury approximates a collection of average people, and that their opinions are the opinions of the "average" person in the community (Brigman 1978). Even if the pool from which jurors are initially drawn is representative of the community, there is no guarantee that, once a jury is ultimately selected for participation in the trial, it is even representative of the jury pool, much less the community, due to attorney challenges and the small sample size (12 jurors). In contrast, social science holds that, to validly determine anything about the collective opinion in the community, it is necessary to gather a sample of people from the community who are selected in such a way as to assure that they are representative of the population from which they are selected (in practice this means that all members of the population have an equal chance of being included in the sample).

The law also seems to presume that there is such a thing as an "average" person—an individual who is average on each of a set of major social dimensions or at least a person of "average" sexual instincts. (Neither the "average person" nor "contemporary community standards" may include a consideration of children in their calculus, [*Pinkus v. United States* 1978]). In essence, this idea is similar to the now discredited concept of the "average man" advanced by the social statistician Quetelet in the nineteenth century. Quetelet suggested that the means of various traits could be combined to form one paradigmatic human being who would represent the "type" for a group, a city, or a nation. Later debate focused on the possibility that such a paradigmatic individual could actually be located in a population (Kruskal and Tanner 1978). The problem with this notion is that while it is possible to find an individual who is average on a few important traits, the probability of finding a single individual who is average on a larger number of social characteristics than two or three is very small.

---

[9] Elements 6 and 7 of obscenity law are not discussed here for several reasons. Element 6, sexual conduct defined by state law, is primarily a matter for legislatures to decide, and social scientists have, in our opinion, little to say about this issue. Element 7, whether the work lacks serious literary, artistic, political, or scientific value, does not involve estimating the opinions of the community (community standards) based on survey research or a film-viewing study such as the one we have described here, and, as such, this form of expert testimony will not be discussed here.

Social scientists only speak of the "average" person as a set of characteristics defining a sample of persons (for example, we reported statistics that allow us to say that the "average" resident of Mecklenburg County is a married white female; who has resided there for 19 years, 3.3 months; and who is a Democrat). But the average person is purely hypothetical. There is no one individual who possesses exactly these characteristics in the sample. Rather than searching for the average person's opinion, social scientists seek the "average" level of agreement or disagreement to questions put to respondents, once a representative sample of persons has been obtained. In this respect, social science may come as close as is possible to the legal concept of the "average person" as a composite or synthesis of the community (*United States v. Treatmen* 1975; *Urbana ex rel. Newlin v. Downing* 1989). In this study, we gathered a cross-section of the community of Mecklenburg County through scientific sampling and computed the "average" level of prurient appeal and patent offensiveness of certain sexually explicit materials among members of the community.

DEFINING CONTEMPORARY COMMUNITY STANDARDS

The second element, contemporary community standards, can also be viewed from either a legal perspective or competing social science perspective. Obscenity cases are unique criminal proceedings in that they are virtually the only cases where a jury is asked to take into account the consensus of the community before making a judgment about the guilt or innocence of a defendant charged with a crime. Jurors are faced, by law, with the burden of articulating community standards, and not even local or state legislative bodies are allowed to usurp the function of the jury by setting guidelines for those standards (*Smith v. United States* 1977). Judges do not always instruct jurors regarding the operating legal definition of *community,* nor will the judge necessarily instruct jurors how to arrive at a consensus on community standards.

Because the jury is called upon to make what amounts to a sociological and psychological assessment of the standards of the community, it would seem imperative that it have access to reliable social-scientific data collected from the community. In U.S. society, when a politician, a news organization, a business, and, particularly, the government itself wants to know how the people in a particular community feel about a particular issue, they rely upon scientifically accepted means of determining these opinions—the selection of a representative cross-section of persons from the community who are surveyed about their opinions. Such evidence serves a similar function in an obscenity case (Rudolf and Fargo 1988).[10]

Expert testimony from scientific studies may not, however, override the jury's final determination of these standards; the opinions of experts being nonbinding on the trier of fact, who may give the expert testimony whatever weight is deemed fit or may ignore it altogether. However, the methods used to collect these data, the data itself, and the means whereby social scientists arrive at an opinion of what materials the "average" person, applying contemporary community standards, would find patently offensive or appealing to the prurient interest should be provided to the jury in order that it may thoroughly execute its charge in an obscenity case. To exclude such evidence and testimony would seem to ensure a guessing game on the part of the trier of fact as to the applicable community standards or, in the words of Mr. Justice Frankfurter, result in a ". . . hit-or-miss subjective view of what [such standards] are believed to be."

Patent Offensiveness. At least two types of "tolerance" can be measured through a community survey—the opinions of a representative sample of persons about what they *think* the *community* tolerates, and the collective opinions of these same individuals about what they *personally* should be allowed to view. In this study, we asked subjects about both. The findings reported here (see table 4) indicate that a much lower percentage of people think the hypothetical "community" tolerates the films they just viewed than is the case when people are asked to report on what they personally tolerate (what they should be allowed to view). If we define community tolerance as the sum of individual responses about what people personally would tolerate, we would conclude that the community of Mecklenburg County is *overwhelmingly* tolerant of these materials, by a margin of 3 to 1.

---

[10] As Rudolf and Fargo point out, before conducting such a survey, care has to be given to the way questions are worded. A number of courts have refused to allow survey results into evidence on the grounds that the questions asked were vaguely or improperly worded (*United States v. Pryba* 1987; *State of North Carolina v. Mayes* 1987). Before asking people whether they believe material depicting nudity and sex are tolerated (or acceptable) or not, one must properly define the terms *nudity* and *sex*. Surveys have been held inadmissible on the grounds that the definitions were either improperly placed in the questionnaire or that the terms did not convey the impact of the visual image depicted in the material. In the present study, we were concerned that subjects, when they responded to the pre-viewing questionnaire, had in mind the types of graphic depictions of sexual conduct that were contained in the film that they subsequently saw. The possibility existed that the subjects' perceptions, based on the descriptions provided before they viewed the films, were different in content or character than the actual material. This concern also went directly to the adequacy of the description of the content and character of sexually explicit material presented in the telephone survey administered in 1987. In order to assess the level of congruence between the definition given to subjects over the phone and their perceptions of the material after viewing it, we asked subjects if the movie they had viewed was what they expected to see, given the definition that had been provided in the prefilm questionnaire (which was identical to the definition given in the 1987 telephone survey). When we considered subjects' responses across all of the films, 88% indicated that the materials viewed were, in fact, what they had in mind when responding to the prefilm questionnaire. Each film groups' responses mirrored these results. Thus, we can conclude that the description used in the prefilm questionnaire, as well as in the 1987 telephone survey, was sufficient to apprise the interviewees of the nature of the charged materials, and the studies do not, therefore, suffer from the problem noted by the *Pryba* court.

In addition, 75% of the subjects indicated that they had viewed a film like the one they had "seen today" sometime before. This percentage is somewhat higher than estimates available from national opinion polls. Yankelovich, Clancy, and Schulman (1986), for example, report that 62% of Americans had seen an X-rated movie at least once. Such national estimates are, in our opinion, highly suspect because respondents may have widely variant definitions of what constitutes an "X-rated" film without a detailed description of the acts portrayed in these films—a procedure not commonly employed in national surveys.

A similar discrepancy between personal standards and perceived community standards has been detected in other investigations of tolerance of sexually explicit materials. Herrman and Bordner (1983) found little consistency between personal standards and the perceptions of community standards in a sample of 750 residents of Atlanta, Georgia, and surrounding counties when respondents were presented with a written list of acts often featured in pornography. Personal judgments were much less conservative than perceived community standards for evaluating the descriptions. The authors noted that, with the exception of rape, little overlap occurred between the two sets of attitudes.

The belief in empirically false propositions about others in the social world found in our study and in other investigations of tolerance for sexually explicit materials may be an example of "pluralistic ignorance" (O'Gorman 1988). The literature on social perception contains an abundance of evidence of pluralistic ignorance, but the direction of the discrepancy is usually in favor of the individual's own attitudes. Usually, people are motivated to overestimate *support* for their values among others (e.g., Fields and Schuman 1976–77)—an effect known as the "egocentric bias." In the present study, we found individuals perceiving less, rather than more, support for their beliefs about sexually explicit materials.

Through what processes might the social environment have misinformed its inhabitants about attitudes toward sexually explicit materials? We propose that the discrepancy between perceptions of the community standard, and the individual adults' own standards in the present study, is the result of recent legal events in Mecklenburg County. During the past two years, several arrests for obscenity and obscenity trials where defendants had been found guilty had taken place in the community. These were well publicized. Subjects may have assumed that, since there had been a large number of arrests and several guilty verdicts, most people in the community must not tolerate these types of materials (a phenomenon we would call "prosecution-induced intolerance"). The greater the attention given law enforcement activities by the media in a community, the more the average observer may assume that citizens of the community are intolerant. However, when members of the community are individually questioned, they may express a much higher level of tolerance for sexually explicit materials.

This misperception, primarily gained from the mass media, may also have consequences for interpersonal interactions, which in turn also influence perceptions about community beliefs. The erroneous belief in lack of tolerance for sexually explicit materials in the community may lead people to be hesitant to speak out honestly about their own opinions for fear they are deviant. Here, another theory of public opinion may come into play. This unwillingness to speak out may be an example of the more general tendency toward a "spiral of silence" in public opinion whereby a silent majority falsely perceives itself to hold a minority opinion and, thus, remains quiet in order to avoid personal ridicule (Noelle-Neumann 1974 and 1984). Mosher (1989) has speculated that public opinion toward pornography and obscenity may be very susceptible to this effect. Most Americans, according to this view, favor the continued availability of sexually explicit material for adults, yet they are very reluctant to either state that they personally enjoy viewing these materials or defend others' rights to view them.

The end result, should these mass communication/social psychological processes be operating, is that the legal system may be an unwitting but critical contributor to the very standard it is trying to discover through the criminal fact-finding process. Members of the community (including jurors themselves) may assume the community does not tolerate sexually explicit materials because such materials are not tolerated by local law enforcement officials (as evidenced by continued prosecutions).[11]

It might be informative to jurors for a social scientist to present evidence about *both* perceptions of the community and individual levels of tolerance. It should be acknowledged, however, that most policy decision making based on public opinion research relies on the summation of individual opinions rather than "guesses" by respondents about others' opinions and behaviors. As social scientists, we place far more confidence in a procedure whereby we ask people if they should personally be free to engage in a specific behavior (such as watching a movie) and summing these individual responses in order to obtain "the average" response, than we would be in asking people to report on what they *perceive* the "average" person would do in the same situation.

---

[11] Whatever the cause, the discrepancies noted in this study between what people perceive is tolerated in their community with what is actually tolerated (as determined by the summation of individual personal opinion responses) have one additional legal implication. These findings call into sharp question the general legal assumption that a judge or jury, without any assistance whatsoever, knows or can determine beyond a reasonable doubt the contemporary community standards by which both prurient appeal and patent offensiveness are to be judged.

Appeal to Prurient Interest. The other element of the obscenity test evaluated in this study is whether the work, taken as a whole, appeals to prurient interest. In our study we found a statistically significant shift toward rejection of the proposition that the material appealed to a morbid, unhealthy, or shameful interest in sex, once subjects had viewed the materials. Since the shift brought subjects back to the baseline levels established in the 1987 telephone survey of Mecklenburg County, it may be that the film/magazine-viewing subjects' initial opinions about the degree to which sexually explicit materials appealed to a "bad," or unhealthy, interest in sex were also affected by "prosecution-induced intolerance." Actually viewing the materials at issue reestablished prevailing community norms. Viewers were "reassured," having now seen the materials for themselves, that they did not appeal to a prurient interest. Our analysis of the mean levels of opinion shift for persons who report having had experience with these materials versus those who have not is consistent with this conclusion. Those who had reported "not seeing films like these before" substantially modified their positions toward less belief in prurient appeal, while those that reported seeing them shifted only slightly or not at all.

Summary and Conclusions

Our findings suggest that the procedures described in the present study, whereby respondents view the specific materials charged in obscenity trials rather than make hypothetical judgments about obscenity, will produce substantially different and, we would argue, more accurate assessments of community standards regarding sexually explicit materials than can be obtained by survey or telephone interviews alone. There may exist a substantial portion of the community who have inaccurate (or nonexistent) perceptions of these materials, before actually seeing what these materials contain. Once they are faced with the materials, our findings suggest that their evaluations may be altered substantially.

Future research should be directed to investigating the robustness of the shift in opinion detected in the present study, to the differences in subjects' perception of the "community" standard relative to their individual standards, collectively measured, in other jurisdictions, and to the magnitude of such changes as a function of the degree of obscenity litigation across jurisdictions. If these shifts and differences can be found elsewhere, they would point to a serious bias in lay and jury decision making whereby laypersons consistently overestimate intolerance of sexually explicit materials in their community.

Appendix

Recently, a number of advances have been made in statistical methods for the diagnosis of, and correction for, subject selection bias (Berk 1983; Berk and Ray 1982). Two of the principal approaches to the correction for sample selection bias are represented by the *selection modeling* methods of Heckman (1979) and the Bayesian *mixture modeling* methods of Rubin (1977). Heuristically, the selection approach models the probability of being a nonrespondent (nonparticipant in the present study) as a function of known characteristics of nonrespondents (e.g., sociodemographic variables) and then enters a function of this probability (called a hazard rate for nonresponse) into a statistical model for the correction of the frequency distribution on a dependent variable (such as one of the questions in table 4) that has been obtained from respondents (participants in the present study). By contrast, the mixture approach takes as given the distribution on a dependent variable for the sample of respondents, constructs statistical models of the unobserved distribution of the dependent variable for nonrespondents, and then mixes the observed and unobserved distributions of the dependent variable and outputs this *mixed* distribution (or some function thereof, such as the mean). Recent methodological work (Wainer 1986) suggests that each approach has its strengths and weaknesses, but there currently is no clear consensus among statisticians that one is always preferable to the other.

We report the results of applying the mixture modeling approach to ascertain the extent to which the "after film" response distributions in table 4 could be biased by sample self-selection. This decision is based largely on the fact that the mixture modeling produces probability intervals as output. These achieve our goal of displaying the sensitivity of the "after film" response distributions to various assumptions about selectivity bias. By comparison, the selection modeling approach focuses on providing an answer to the question of whether or not significant selection bias exists under a single, untestable assumption. Nonetheless, it should be noted that we also subjected the data on film-viewing responses in table 4 to analyses by the selection modeling procedure. Results from the latter analyses are consistent with those from the mixture modeling analyses reported here.

The Rubin (1977) method can be thought of as summarizing the results of simulations in which one uses data on both independent and dependent variables for respondents to generate reasonable hypothetical responses for nonrespondents. The method has two key parameters. The first parameter, θ_1, formalizes subjective notions about how similar the regression slopes relating a dependent response variable Y to a set X of independent variables are likely to be for respondents and nonrespondents. Technically, θ_1 is the subjective coefficient of variation (the standard deviation divided by the mean) for the nonrespondents' regression coefficients. It is used to place an interval (e.g., a 95% or 2-standard-deviation-relative-to-estimated-coefficient interval) around the regression coefficients for the respondents within which the analyst expects the regression coefficients for the nonrespondents, if it were possible to estimate them, would fall. Similarly, the second parameter, θ_2, formalizes subjective notions about how similar the analyst believes the expected value (mean) of Y is likely to be for respondents and nonrespondents with equal X means (i.e., with equal means on those independent variables that are observed for both groups). In practice, the mixture modeling method is applied with θ_1 and θ_2 taking on a range of numerical values in order to study the sensitivity of the bias correction outputs to these parameters. Rubin (1977) indicates that coefficients of variation are used because they are well suited to summarizing notions of large and small effects (e.g., $\theta_1 = 0.01$ seems small, while $\theta_1 = 1.0$ seems large). The method assumes unbiased (i.e., centered at the expected values of the regression slopes for respondents) and normal prior distributions of the regression and subjective parameters and that certain of these distributions are joint while others are independent. The sensitivity of numerical results to these assumptions is not known, but, based on the behavior of similar statistical methods, it is reasonable to expect that the sensitivity declines as the sample size increases.

In the present context, the dependent Y variables to which we applied Rubin's mixture modeling selection bias correction procedure are the three "after film-viewing" response variables (table 4), while X is a matrix containing observations on the eight independent sociodemographic variables on which information was obtained from all adults who completed the telephone survey. (Note that data on both the three, Y-response variables and the eight, X-variables are available for the 109 telephone survey respondents who participated in the film-viewing study *and* viewed one of the X-rated films, whereas only data on the eight independent X-variables are available for the 100 telephone survey respondents who declined to participate after being informed that some films would be X-rated.)

References

Bantam Books Inc. v. Sullivan. 1963. 372 U.S. 58, 83 S.Ct. 631, 9 L.Ed. 2d 584.

Beckett, J. Steven, and Roderick A. Bell. 1979. "Community Standards: Admitting a Public Opinion Poll in an Obscenity Case." *Case and Comment* 84(2): 18–24.

Berk, Richard A. 1983. "An Introduction to Sample Selection Bias in Sociological Data." *American Sociological Review* 48:386–98.

Berk, Richard A., and Subhas C. Ray. 1982. "Selection Bias in Sociological Data." *Social Science Research* 11:301–40.

Brockett v. Spokane Arcades Inc. 1985. 472 U.S. 491, 105 S.Ct. 2794, L.Ed. 2d 394.

Brigman, William E. 1978. "The Controversial Role of the Expert in Obscenity Litigation." *Capital University Law Review* 7:519–51.

Converse, Jean M., and Stanley Presser. 1986. *Survey Questions: Handcrafting the Standardized Questionnaire*. Beverly Hills, CA: Sage.

Donnerstein, Edward I., and Ben Bates. 1987. "Public Opinion and the Two Pornography Commissions: Comparing the Incomparable." *Southwestern Mass Communication Journal* 6(1): 1–10.

Festinger, Leon. 1957. *A Theory of Cognitive Dissonance*. Palo Alto, CA: Stanford University Press.

Fields, James M., and Howard Schuman. 1976–77. "Public Beliefs about the Beliefs of the Public." *Public Opinion Quarterly* 40:427–48.

Glassman, Marc B. 1978. "Community Standards of Patent Offensiveness: Public Opinion Data and Obscenity Law." *Public Opinion Quarterly* 42:161–70.

Hamling v. United States. 1974. 418 U.S. 87, 94 S.Ct. 2887, 41 L.Ed. 2d 590.

Heckman, James J. 1979. "Sample Selection Bias as Specification Error." *Econometrica* 45:153–61.

Herrman, Margaret S., and Diane C. Bordner. 1983. "Attitudes toward Pornography in a Southern Community." *Criminology* 21:349–74.

Kruskal, William H., and Judith M. Tanner. 1978. "A. J. Quetelet." In *International Encyclopedia of Statistics*. New York: Free Press.

Kaplan v. California. 1973. 413 U.S. 115, 37 L.Ed. 2d 492.

Lamont, J. 1973. "Public Opinion Polls and Survey Evidence in Obscenity Cases." *Criminal Law Quarterly* 15:135.

Miller v. California. 1973. 413 U.S. 15, 93 S.Ct. 2607, 37 L.Ed. 2d 419.

Mosher, Donald L. 1989. "Threat to Sexual Freedom: Moralistic Intolerance Instills a Spiral of Silence." *Journal of Sex Research* 26:492–509.

Mueller, John. 1988. "Trends in Political Tolerance." *Public Opinion Quarterly* 52:1–25.

Noelle-Neumann, Elisabeth. 1974. "Spiral of Silence: A Theory of Public Opinion." *Journal of Communication* 24:43–51.

Noelle-Neumann, Elisabeth. 1984. *The Spiral of Silence: Public Opinion—Our Social Skin*. Chicago: University of Chicago Press.

O'Gorman, Hubert J. 1988. "Pluralistic Ignorance and Reference Groups: The Case of Ingroup Ignorance." In *Surveying Social Life: Papers in Honor of Herbert H. Hyman*, ed. H. J. O'Gorman. Middletown, CT: Wesleyan University Press.

Osborne v. Ohio. 1990. 110 S.Ct. 1691, 109 L.Ed. 2d 98.

Paris Adult Theatre I v. Slaton. 1973. 413 U.S. 49, 93 S.Ct. 2628, 37 L.Ed. 2d 446.

Pinkus v. United States. 1978. 436 U.S. 293, 98 S.Ct. 1808, 56 L.Ed. 2d 293.

Pope v. Illinois. 1987. 481 U.S. 497, 107 S.Ct. 1918, 95 L.Ed. 2d 439.

Rubin, Donald R. 1977. "Formalizing Subjection Notions about the Effect of Nonrespondents in Sample Surveys." *Journal of the American Statistical Association* 72:538–43.

Roth v. United States. 1957. 354 U.S. 476, 77 S.Ct. 1304, 1 L.Ed. 2d 1498.

Rudolf, David S., and Marjorie Fargo. 1988. "The Obscenity Trial in Rural America: Preparing to Win." *The Champion* (November): 5–16.

Saliba v. State of Indiana. 1985. 475 N.E. 2d (Ind. App. 2d) 1181.

Schauer, Frederick. 1976. *The Law of Obscenity*. Washington, DC: Bureau of National Affairs.

Smith, Tom W. 1987. "The Use of Public Opinion Data by the Attorney General's Commission on Pornography." *Public Opinion Quarterly* 51:249–67.

Smith v. California. 1959. 361 U.S. 147, 80 S.Ct. 215, 4 L.Ed. 2d 205.

Smith v. United States. 1977. 431 U.S. 291, 97 S.Ct. 1756, 52 L.Ed. 2d 324.

Stanley v. Georgia. 1969. 394 U.S. 557, 89 S.Ct. 1243, 22 L.Ed. 2d 542.

State of North Carolina v. Anderson. 1988. 322 N.C. 22, 366 S.E. 2d 459, *cert. denied*, 109 S.Ct. 513 (1988).

State of North Carolina v. Cinema Blue of Charlotte, Inc. 1989. State of North Carolina, County of Mecklenburg, Docket nos. 88 CR 546966 et al.

State of North Carolina v. Mayes. 1987. 323 N.C. 159, 371 S.E. 2d 476, *cert. denied*, 109 S.Ct. 792 (1989).

United States v. Pryba. 1987. 678 F.Supp. 1225, *aff'd.*, 900 F.2d 748, *cert. denied*, 111 S.Ct. 305.

United States v. Treatmen. 1975. 524 F.2d (8th Cir.) 320.

United States v. 2,200 Paperback Books. 1977. 565 F.2d (9th Cir.) 566.

Urbana ex rel. Newlin v. Downing. 1989. 43 Ohio St. 3d 109, 539 N.E. 2d 140.

Wainer, Howard, ed. 1986. *Drawing Inferences from Self-Selected Samples*. New York: Springer-Verlag.

Yankelovich, Clancy, and Schulman. 1986. "Pornography: A Poll." *Time*, July 21, 22.

FAMILY LIFE AND AMERICAN POLITICS: THE "MARRIAGE GAP" RECONSIDERED

ERIC PLUTZER and
MICHAEL McBURNETT

Abstract In recent years, several writers have identified marital status as a potentially important line of political cleavage, observing that singles are more likely to vote Democratic than married voters are. Changes in both the structure of American families and in the salience of "family politics" in the policy arena suggest increased attention to the political consequences of marital status and family life-style. This paper contributes to advancing theory concerning the relationship between family life and politics, and empirically evaluating several competing hypotheses concerning the so-called marriage gap in the 1972 through 1988 presidential elections.

Introduction

In recent years several writers have identified marital status as a potentially important line of political cleavage (Gerson 1987; Kingston and Finkel 1987; Plissner 1983; Weisberg 1987). These researchers, as well as several journalists (Clymer 1983; Dowd 1984; New York Times 1984), have observed that singles are more likely to vote Democratic than married voters are. Because some have suggested that the "marriage gap" is likely to increase in the future (Gerson 1987), and since the percentage of singles is projected to increase as well (Kingston and

ERIC PLUTZER is Assistant Professor of Political Science at Iowa State University. MICHAEL MCBURNETT is Assistant Professor of Political Science at the University of Illinois. This research was supported, in part, by a training grant to the first author from the National Institute of Mental Health (T32 MH 15789). The data were originally collected by the Institute for Social Research and made available through the Inter-University Consortium for Political and Social Research. The authors appreciate the comments of Doug Downey, Liane Kosaki, Dianne Schmidt, and Herbert Weisberg. The authors, of course, are responsible for any errors or omissions that remain. Address all correspondence to the first author at the Department of Political Science, 503 Ross Hall, Iowa State University, Ames, IA 50011-1204.

Public Opinion Quarterly 55:113–127 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-01/$2.50

Finkel 1987), the political consequence of marital status looms as a significant unexplored issue in political science.

Two empirical papers that employed data from the National Election Studies to explore the marriage gap in a more rigorous manner were published in 1987 (Kingston and Finkel 1987; Weisberg 1987). Unfortunately, these two papers arrived at rather different conclusions. In this paper, we dissect both the methods and rationale employed in previous studies in an attempt to reconcile their contradictory findings. For those elections in which we find evidence for a marriage gap, we attempt to evaluate all competing explanations, including one that locates the problem of the marriage gap as a particular instance of the relationship between family life and politics.

The paper proceeds as follows. First, we summarize the empirical evidence for the existence of the marriage gap. Second, we review various attempts to explain its existence. Third, we summarize some recent perspectives from the women-and-politics perspective and infer a theoretical basis for making more precise distinctions among those who are married, as well as between married and singles. This "life-style politics" approach contends that political cleavages may be based on life-style differences (Page and Clelland 1978; Plutzer 1991; Wildavsky 1987) between those in "traditional" and those in "nontraditional" family arrangements. Each perspective implies somewhat different hypotheses regarding the presidential vote; we elaborate and test these using data from the National Election Studies (NES).

The Marriage Gap in American Politics

In examining exit polls from the 1982 congressional elections, Plissner (1983) observed that single voters favored Democrats by 26 percentage points, while married voters favored Democrats by a mere 52 to 48 margin. Gerson (1987) also examined exit polls and reports a marriage gap of 12% for the 1984 presidential elections.

Weisberg (1987) examined NES data for each presidential election since 1964 and found a sizable gap between married and nonmarried voters in every presidential election since 1972. After controlling for race and income, however, Weisberg reports that the marriage gap disappeared in all years except 1972.

Kingston and Finkel (1987) used the 1984 NES study to examine differences among married, single, and "previously married" respondents. After controlling for a variety of background factors, Kingston and Finkel report significant net differences between married and never married voters with respect to the presidential vote.

The discrepancy between the findings of Kingston and Finkel (significant net effects) and Weisberg (apparently spurious correlation) in the 1984 election is especially problematic and serves as the major motivation of this paper. Both used the 1984 presidential vote as a dependent variable and both used the same NES data. Kingston and Finkel employed the same control variables as Weisberg, plus several others (including age, sex, education, and place of residence), and still found statistically significant net differences. One possible explanation is that the additional independent variables used by Kingston and Finkel have brought out a suppressed effect (for a discussion of statistical suppression, see Rosenberg 1968). We suspect, however, that the difference is due to the different definitions of marital status: the relatively crude dichotomy used by Weisberg in comparison with the (slightly less crude) trichotomy employed by Kingston and Finkel. We explore both possibilities in this paper.

EXPLANATIONS

Weisberg (1987) offers a "demographic" explanation for the marriage gap. Because blacks are both more likely to vote Democratic and less likely to be married, the zero-order marriage-vote relationship may be due to the unmeasured effects of race. Also, since married couples tend to have higher family incomes than singles, and family income is positively associated with Republican voting, the zero-order association may be due to the unmeasured effects of income. Weisberg supported this line of argument with the fact that the marriage gap disappeared after controlling for race and family income in three of four elections (see his table 3, col. 4).

Plissner speculates that "married people are more likely to own property and to worry about protecting it, . . . to have, or to expect, children and, if so, to take a benign view of authority and a dim view of social disorder" (1983, 53). Thus, he concludes, married individuals may be more likely to respond to the conservative values enunciated by Republicans. It would follow, then, that property ownership and/or the presence of children would act as intervening variables between marital status and political orientation. These hypotheses have yet to be systematically tested.

Kathleen Gerson (1987) offers a somewhat different explanation for the observed political differences between married and single voters. Her explanation concerns women voters only, although we argue below that it may be extended to men as well. Gerson argues that the fundamental political division among women is between those who have a commitment toward traditional domestic responsibilities and those whose commitments lie in nondomestic realms such as paid employment. These commitments develop through the life course as a consequence of such factors as career opportunities and the nature of women's relationships with men. A direct quotation should make Gerson's line of argument clear.

> Unanticipated encounters with changing work and family structures led some women to veer toward domesticity and others to veer away from it. Those women who experienced instability in their relationships with men, who encountered often unanticipated work opportunities, and who experienced severe economic squeezes in their households tended to develop strong work commitments. . . . (1987, 215)

Whatever commitments women find themselves with, Gerson claims, they will tend to adopt ideologies and support candidates and policies that support life-styles they are committed to (cf. Wildavsky 1987). Speaking of women who develop nondomestic commitments, Gerson argues that they seek "to gain structural and ideological acceptance for gender equality" (1987, 216).

On the other hand, women who have opted for traditional marriages have seen the ideological and structural supports of that life-style diminished.

> In the face of declining marital stability, an eroding male paycheck, and the increasing devaluation of full-time motherhood, they must struggle to defend traditional ideological and structural supports that their mothers took for granted only a few decades ago. (1987, 217)

Though Gerson's hypotheses are based on her research of women, they provide the basis for a more comprehensive theory of family life and politics. Gerson's approach is loosely based on the assumption that political conflicts concerning social, as opposed to economic, issues can be understood as a *defense of life-style* (Gusfield 1963; Page and Clelland 1978; Wildavsky 1987; Wood and Hughes 1984). Like their female counterparts, men in traditional marriages may have interests in both ideological and material supports for such family lifestyles. Similarly, men whose wives are employed have interests in such material goals as equal pay for women and affordable day-care. Men whose wives work also have an interest in ideologies that legitimate this more modern family arrangement. We find some evidence for this line of argument provided in Klein's analysis of the 1972 NES data showing that men as well as women differ in their support of feminism according to the traditionality of their household (Klein 1984, table 7.1). Thus, we expect a division among married men as well as among married women. In this case, however, the focus is not on men's labor-force experience but that of their wives.[1]

[1] Because we draw upon theories from the fields of gender roles and women and politics, it is obvious to wonder whether the analysis should be undertaken separately for men and women. We find little in the literature that leads to hypotheses predicting

THE QUESTION OF DIVORCE AND WIDOWHOOD

Previous studies have tended to either ignore the divorced and widowed, combine them with other "singles," or group them together in a category called "previously married." Although we have no body of established theory to rely on, we suspect that these two groups should *not* be combined. We tentatively purpose a few hypotheses as a logical extension of the reasoning presented thus far.

Divorce. Divorce represents, in most cases, the failure of "traditional family life." When men and women choose to end a marriage, they implicitly are opting for a less traditional life-style (Ehrenreich 1983; Klein 1984). It seems reasonable to suspect that those who have lower levels of commitment to "traditional family life" will be both more liberal and more likely to get divorced. When choosing divorce for reasons of necessity (e.g., to escape a violent marriage) or when divorced by one's spouse, one nevertheless finds oneself faced with having to fulfill nontraditional roles. So it is also possible that once divorced—for whatever reason—individuals have fewer interests served by the party that is more aligned with traditional family ideology.

Widowhood. For reasons just mentioned, it seems inappropriate to combine the widowed with the divorced. But we have little to go on in predicting the political orientation of widows. There is reason to suspect that those who are somewhat conservative on family issues are more likely to become widows. After all, nontraditionalists are more likely to have never married in the first place and, if so, are more likely to have been divorced. Thus, widowhood may be seen, in many cases, as the culmination of a lifetime of traditional family life. On the other hand, widowhood for men and women may place them—like the divorced—in situations where they must broach traditional role boundaries (e.g., widowers learning domestic skills, widows being forced to become the principal wage earner). So we make no prediction about where the widowed will fall in the partisan spectrum.

Data and Methods

We employ the National Election Studies for presidential elections from 1972 through 1988 to explore the relationship between family life and politics. The NES data are ideal for our purposes for several rea-

---

differences between men and women, but nevertheless explored the possibility empirically. We find no significant differences when we estimate our models separately for male and female voters.

sons. In addition to the high quality and familiarity of the NES data, the fact that they were used by Kingston and Finkel (1987) and Weisberg (1987) in their explorations of the marriage gap permits us to replicate their initial findings and then incrementally test additional hypotheses..Any new conclusions from the present study could only be attributed to differences in model specification, rather than differences in the time or method of data collection. Moreover, one recent empirical work suggests that the effects of family status on feminist ideology increased between 1977 and 1985 (Plutzer 1988); the availability of the 1988 data allow us to extend the series of elections explored and see if the effects of family life have increased in recent years.

DEPENDENT VARIABLE

In this paper, we use the respondent's self-reported presidential vote as the dependent variable. Derived from the NES postelection sample, this variable is scored one for a vote for the Democratic candidate and zero for the Republican.[2] Respondents who did not vote or refused to answer this item are deleted from the analysis.

INDEPENDENT VARIABLES

The independent variables fall into three categories. First are a variety of social and demographic background factors. These include dummy variables for race (nonwhite = 1); southern, western, and midwestern place of residence (east omitted); ordinal variables for family income (17 categories); size of community in which the voter grew up (recoded to three categories consistent with the analysis of Kingston and Finkel); and interval variables for age and years of education.

Marital status will be measured in four ways. First, a simple dichotomy as used by Weisberg in his research (married = 1, all others = 0). Kingston and Finkel's trichotomy is measured with dummy variables for married and previously married respondents (never married is the omitted category). Gerson's suggestion of adding a distinction between traditional and nontraditional married respondents produces a four-category measure: traditional married (wife does not work), nontraditional married (wife employed), and previously married (never married is again omitted). Finally, we have a five-category specification that allows for distinguishing among the previously married respondents along the lines we suggested above: the previously married variable is replaced with a dummy variable for the divorced and one for the widowed.

---

[2] In 1980, John Anderson ran as an independent. We coded presidential vote two ways: (1) Carter vs. Reagan with Anderson voters ignored, and (2) Carter/Anderson vs. Reagan. Both methods led to similar results. We report only the latter analyses in this paper.

Third, we have three measures that we treat as intervening variables. These include a dummy variable indicating home ownership and one indicating whether or not the respondent has any children in the home; these permit evaluating explanations proposed by Plissner (1983). We also include a seven-point party identification measure. All independent variables with missing values are recoded to the mean.

STATISTICAL CONSIDERATIONS

Because of the dichotomous nature of the dependent variable, we estimate our models using logistic regression. This complicates comparisons with previous studies that employed different analytic techniques; Kingston and Finkel (1987) and Weisberg (1987) both used ordinary least squares regression. As a result, we also repeated our analyses using OLS to determine if our results differ due to the different specification of independent variables or better statistical methods; we note in advance that we found no disparities due to the use of OLS.

A second consideration involves the use of a weighted sample. Weisberg weights all respondents by the number of adults in the household. This has the advantage of providing a statistical description that is as close to the electorate as possible. The disadvantage is that the sample size is inflated, making it easier to achieve statistical significance. For the purposes of this paper we elected to use the "smaller," unweighted sample because it provides more conservative tests of our hypotheses.[3]

MODELS TO BE ESTIMATED

We initially estimate five nested logistic regression models for each presidential election, beginning with a baseline model which includes only social and demographic background variables (this model is la-

[3] In a personal communication, Weisberg has noted that weighted analysis has a second virtue. It corrects for the problem that the likelihood of being selected by the CPS/Kish sampling scheme is necessarily correlated with family size and, therefore, with the central dummy variables in this analysis. A third alternative is to weight the sample in proportion to the number of politically eligible adults in a way that retains the original sample size. While the test statistics for individual variables may be based on effective cell sizes that differ slightly from the number of actual subjects, the overall goodness-of-fit statistics (which are very sensitive to sample size) will be based on a more representative distribution and also based on the actual number of subjects. We reestimated all equations using such a proportional weighting scheme and achieved results that were nearly identical to those reported below.

Table 1. Goodness of Fit for Logistic Regression of Presidential Vote Models

		−2*log (Likelihood Ratio)				
Model	*df*	1972	1976	1980	1984	1988
M0 (baseline)	8	1826.14	2203.34	1201.02	1679.20	1497.25
M1 (M0 + married)	9	1821.43^{a}	2203.33	1199.81	1678.69	1497.15
M2 (M0 + married + prev. married)	10	1820.51	2201.88	1198.12	1670.82^{a}	1494.03
M3 (M0 + traditional married + nontrad'l married, prev. married)	11	1817.95^{a}	2198.56	1197.86	1666.99^{b}	1493.99
M4 (M0 + traditional married + nontrad'l married, div., widowed)	12	1817.66	2196.64	1197.79	1666.77	1493.75
N		1,587	1,742	958	1,366	1,195

NOTE: Preferred models are underlined.

[a] Improvement in fit from model M0 is significant at the .05 level.

[b] Improvement in fit from model M0 and the immediately preceding model significant at .05 level.

beled M0 in tables). The second model (M1) includes the dichotomous measure of marital status used by Weisberg, and the next (model M2) uses the trichotomous measure employed by Kingston and Finkel. The next equation (M3) uses an alternative measure of "traditional" or "nontraditional" married men and women as implied by the sociological literature reviewed above; the last model (M4) divides the previously married into divorced and widowed categories. Once we have ascertained the best-fitting model(s), we will add the intervening variables of party identification, home ownership, and number of children in order to better interpret our findings.

Results

Because the first stage in the analysis involved estimating 25 logistic regression equations, we initially report only the summary goodness-of-fit statistics in table 1. The models were estimated by maximum likelihood (using SAS) and, following convention, we report the likelihood ratio chi-square (−2 times the likelihood ratio) that is distributed as chi-square. The likelihood ratio (L^2) decreases as models fit the data better. Since the models are nested and each succeeding model has one additional variable compared to the previous model, a more complex model is preferred to the model immediately preceding it only if the improvement in L^2 is statistically significant across one degree of freedom; it is preferred to the model two steps earlier if the improvement is significant across two degrees of freedom, and so on (those familiar with OLS will note the similarity in logic to evaluating the improvement in R^2 when sets of dummy variables are added to a regression equation).

Looking first at the L^2 goodness-of-fit statistics, we see that, in comparison to the simple baseline model, the addition of a simple dichotomous measure of marital status (model M1) is significant only in 1972; this replicates the results reported by Weisberg (using OLS and weighted data). In 1984, the model employing the three-category definition of marital status fits significantly better than the baseline model. This is reassuring because it reconciles the apparently contradictory findings reported in Kingston and Finkel 1987 and Weisberg 1987; the former's three-category scheme yields the best model *only* in 1984, the only election they analyzed. So far, the baseline model (no marital effects) is the preferred model for 1976, 1980, and 1988, while M1 is preferred in 1972, and M2 in 1984.

Model M3 divides the married into "traditional" and "nontraditional" groups based on the employment status of the wife. In 1972 and 1984 this model fits significantly better than the baseline model, but only in 1984 is it better than the previous preferred model.[4] Model M4 extends the traditional/nontraditional logic further by distinguishing between the divorced and widowed (collapsed into the category "previously married" in M2 and M3). This distinction fails to significantly improve the fit in any of the five elections. So the preferred models are the no effects model (M0) in 1976, 1980, and 1988; the simple "dichotomy" model (M1) is best in 1972; and the "family lifestyle" model (M3) in 1984.

The two elections in which family variables have net effects beyond the impact of the eight background variables are of primary interest. Table 2 reports the estimated logistic regression coefficients for the preferred models in 1972 and 1984. In 1972, the coefficient for being married (contrasted with all unmarried men and women) is $-.297$; taking the antilog (e^b) we estimate that the odds of married respondents voting for McGovern were 26% lower than the odds of the unmarried voting for McGovern. In 1984, the odds of the nontraditional married

---

[4] In fact, the improvement of L^2 of 3.83 has a p-value that rounds to .05 (.0503 to be more precise). Since our hypothesis predicted the direction of the effect, however, the probability of getting an improvement of 3.83 in the predicted direction is, in fact, .025. Thus, we conclude that M3 is the preferred model even though the p-value is not strictly less than .05.

Table 2. Logistic Regression Equations for Preferred Models

	1972 (M1)		1984 (M3)	
Variable	*b*	SE	*b*	SE
Intercept	0.547	.416	0.305	.445
Nonwhite	2.456	.241	1.922	.229
Education	−0.009	.023	−0.021	.025
Family income	−0.047	.017	−0.070	.013
Age	−0.018	.004	0.007	.004
Southern residence	−0.635	.165	−0.075	.178
Western residence	−0.280	.181	0.252	.184
Midwestern	−0.162	.151	−0.095	.173
Community size	0.423	.132	0.292	.094
Married	−0.297	.136	—	—
Nontraditional married	—	—	−0.337	.186
Traditional married	—	—	−0.659	.212
Previously married	—	—	−0.654	.215
N	1,587		1,366	
L^2	1,821.43		1,666.99	

Probability of Voting for McGovern		*Probability of Voting for Mondale*	
Married	.33	Nontraditional married	.44
Not married	.40	Traditional married	.36
		Previously married	.37
		Never married	.53

voting for Mondale were 29% less than the never married, while the odds of both the traditional married and the previously married were 48% less than singles.

Since voting results are rarely reported in terms of odds, we highlight the implications of these coefficients by computing the estimated probability of voting for the Democratic candidate in each year. The effects in logistic regression are not linear and depend on the values of other independent variables. Thus we follow convention and calculate estimated probabilities for each category of family status, evaluated at the mean of all the other variables in the equation. These estimated probabilities are reported below table 2.

In 1972, the "average" married respondent had a probability of voting for McGovern of .33, while unmarried voters voted for McGovern 40% of the time. This estimated marriage gap of 7.0% is comparable to the gap of 5.5% estimated by Weisberg using slightly different methods.

Table 3. Goodness of Fit for Logistic Regression Models with Intervening Variables

Model	*df*	−2*log (Likelihood Ratio) 1972	−2*log (Likelihood Ratio) 1984
$\mathrm{M0}_{p}$ (M0 + party ID)	9	1433.04	988.63
$\mathrm{M1}_{p}$ (M1 + party ID)	10	1427.97[a]	978.12[a]
$\mathrm{M2}_{p}$ (M2 + party ID)	11	1426.88[a]	973.25[b]
$\mathrm{M3}_{p}$ (M3 + party ID)	12	1425.28	970.64[a]
$\mathrm{M4}_{p}$ (M4 + party ID)	13	1425.28	970.37[a]
$\mathrm{M0}_{hcp}$ (M0 + home owner + child + party ID)	11	1427.01	975.20
$\mathrm{M1}_{hcp}$ (M1 + home owner + child + party ID)	12	1423.54	975.13
$\mathrm{M2}_{hcp}$ (M2 + home owner + child + party ID)	13	1422.68	962.91[b]
$\mathrm{M3}_{hcp}$ (M3 + home owner + child + party ID)	14	1421.33	960.17[c]
$\mathrm{M4}_{hcp}$ (M4 + home owner + child + party ID)	15	1421.32	959.63[a]

NOTE: Preferred models are underlined.

[a] Improvement in fit from model M0 is significant at the .05 level.

[b] Improvement in fit from model M0 and immediately preceding model significant at .05 level.

[c] Improvement in fit from model M0 and model two steps preceding significant at .05 level.

In 1984, the estimated probabilities highlight a sizable 8% difference between traditional married and nontraditional married voters, a difference that is hidden by the usual coding of *legal* marital status. The estimated probabilities also highlight the 16% difference between the never married and previously married, which were collapsed into the same category of "unmarried" in Weisberg's analysis of the 1984 election.

At this point, two questions immediately suggest themselves. Why are there family effects of some kind in 1972 and 1984 but not in the two intervening elections? Exactly how does family life influence presidential vote in 1972 and 1984? We deal with the latter question first, delaying the former until our concluding discussion.

HOW DOES FAMILY LIFE AFFECT THE VOTE?

We first explored the question of whether this family effect acts through the intervening variable of party identification. We added party identification and repeated our estimation of nested models. The L^2 statistics are reported in the upper part of table 3; they show that

Weisberg's two-category approach continues to be the best model in 1972; in 1984, the four-category model is no longer significantly better than Kingston and Finkel's three-category approach. The implication is that the distinction between traditional and nontraditional married voters in 1984 acted through party identification in this election, but that the effect of marital status, more narrowly defined, is net of party preference.

We then sought to explore the explanations proposed by Plissner (1983): marital differences in voting for president can be explained by home ownership and the presence of children which, presumably, lead voters to support the more conservative party. Since marital status has net effects after controlling for party, this cannot be correct. But it is possible that home ownership and the presence of children might still lead voters to support the more conservative *candidate*, after controlling for party identification.

To evaluate this possibility, we added dummy variables for home ownership and the presence of children in the household to the models and, again, repeated the estimation of the five nested-models for 1972 and 1984. The summary of these analyses appear in the lower part of table 3. After including the two intervening measures, the distinction between married and unmarried voters in 1972 is no longer significant.[5] In 1984, on the other hand, the inclusion of these two variables fails to diminish the effect of family status.

The Marriage Gap Reconsidered

The marriage gap is real in so far as the term refers to *observed differences* in voting between those in different marital statuses (Weisberg 1987). Observation of this gap led to several different causal explanations and we tried to evaluate each of these in this paper. We had hoped that our analysis, spanning five different elections, would help clarify matters by giving clear support to one of these theories. That is not what happened. In fact, the evidence weighs against all three as *general* explanations of the observed marriage gaps since 1972.

Weisberg suggested that the marriage gap could be explained by the prior effects of race and income (the single exception, 1972, perhaps being attributable to chance). But marital status has some net impact on the vote in two of the five years.[6] Plissner suggested that the effect of marriage acts through the presence of children and home ownership, but in 1984 the effect of marital status was undiminished after controlling for these two intervening variables. Recent work by Gerson suggested that marital status per se was an indicator of traditional versus nontraditional family life-style; but when traditional life-style was measured more directly, it significantly improved prediction of the vote only in 1984.[7]

---

[5] Having found that the marital effect disappears after adding the two intervening variables, we were curious about whether the crucial variable was home ownership or children. So we estimated the models with these added alone. Neither, by itself, caused the marital effect to disappear.

The central problem in the current analysis is why family variables are net predictors of voting in some elections and not others. Admittedly, speculation about the marriage gap is recent, and the theoretical explanations are not complex. The various explanations proved useful in guiding our analyses but all were found wanting.

We suspect that more analyses over a longer span of time will be necessary to better understand the role of family life in presidential elections. There are several reasons why.

First, with "nontraditional marriages" becoming more common, the distinction between different types of married life-styles could become more pronounced; with more wives working full time, the interests of nontraditional families may become increasingly salient and crystallized in a way that encourages interest-based voting behavior. Conversely, as economic trends make it more difficult to maintain families in which the husband is the sole source of income, those in traditional families may feel their way of life under attack and may respond to candidates and parties which lament the "breakdown of the American family" and give moral justification for the sacrifices necessary to maintain traditional family roles.

This suggests an additional possibility: it is possible that voters' family life-style interacts with the particular choice of candidates offered to the electorate in a given election. When the two candidates are similar (or are perceived as similar) on issue positions relevant to family life, we should not expect voters to divide along lines of family status. If this is so, it might explain the failure to find any net marital

[6] Of course, chance is always a potential explanation. According to traditional significance tests (bearing in mind the assumptions underlying them), the likelihood of two significant effects in five tries is slim. Moreover, the logistic regression coefficients for married status in 1972 and 1984 are far larger than those estimated for marital status in each of the other years: .008 in 1976, −.17 in 1980, and −.04 in 1988.

[7] In a personal communication, Gerson suggests that it is not surprising that the type of coding employed in this analysis does not produce stronger effects. The methods employed here, however crude, are consistent with the argument advanced in her 1987 article. But the more detailed treatment in her book (1985) suggests that changes in family status, domestic responsibilities, and relationships influence ideology over time. Thus, the snapshot categorization employed here is not a fair test of her theoretical contribution.

effects in 1976 and 1980 when the Democrats nominated Jimmy Carter, a social moderate and self-identified "born again Christian."[8] Unfortunately, this line of reasoning would lead to the expectation of a large family effect in 1988, when voters perceived a substantial difference in the candidates' positions toward women's roles.

The effects of family status in 1972 and 1984 were sufficiently large to have had an impact in close elections. It is possible these effects occurred by chance but not very likely; so it is difficult to simply dismiss them. On the other hand, the limited number of elections we are able to analyze do not permit us to offer a good explanation for why family effects were observed in those particular years but not others. Thus we are inevitably drawn to a final possibility, one that is unsatisfying to those who hope that the study of voting behavior would become a cumulative science, increasingly characterized by valid generalizations. We must conclude that sizable family effects are *possible* but emerge only under particular combinations of circumstance—combinations we are unable to identify at this time.

For now, the marriage gap identified by Kingston and Finkel in the 1984 election and by Weisberg in the 1972 election remain curiosities. They are curiosities we should nevertheless keep our eye on. Family life continues to undergo change in the United States. Policies intended to facilitate nontraditional family life-styles are increasingly a part of the national agenda as evidenced by national child-care policies recently proposed by both parties. In addition to increased attention from political elites and the parties, family issues are currently high priority goals of the women's movement. As Weisberg (1987, 342) concludes, family-based differences in voting (whatever the causes) may generate political campaigns that appeal more directly to family-based segments of the electorate. These societal and political changes will have consequences for mass political behavior; the research agendas of political scientists must stay abreast of these changes and we hope this paper stimulates new inquiries into the role of the family in American political life.

---

[8] Unfortunately, it is not possible to assess this directly. A very crude evaluation, however, can be made by examining the NES sex roles scale questions. These items ask respondents to rate themselves and the candidates on a seven-point scale ranging from "women should have an equal role with men running business, industry, and government" to "women's place is in the home." Unfortunately, the candidate rating questions were not asked in 1984—the year with the greatest family effects on voting. Voters perceived McGovern as .55 points more liberal than Nixon, Carter .22 points more liberal than Ford, Carter .63 points more liberal than Reagan, and Dukakis 1.1 points more liberal than Bush. To the extent that this item is a valid proxy, these results help explain the absence of a marriage effect in 1976 but not in 1980 or 1988.

References

Clymer, Adam. 1983. "Poll Shows a Marriage-Single Gap in Last Election." *New York Times*, January 6.

Dowd, Maureen. 1984. "Singles Show Less Support for Reagan." *New York Times*, December 16.

Ehrenreich, Barbara. 1983. *The Hearts of Men: American Dreams and the Flight from Commitment*. Garden City, NY: Anchor Press.

Gerson, Kathleen. 1985. *Hard Choices*. Berkeley: University of California Press.

Gerson, Kathleen. 1987. "Emerging Social Divisions among Women: Implications for Welfare State Policies." *Politics and Society* 15(2): 213–21.

Gusfield, Joseph. 1963. *Symbolic Crusade*. Urbana: University of Illinois Press.

Kingston, Paul William, and Steven E. Finkel. 1987. "Is There a Marriage Gap in Politics?" *Journal of Marriage and the Family* 49(February): 57–64.

Klein, Ethel. 1984. *Gender Politics*. Cambridge: Harvard University Press.

New York Times. 1984. "Portrait of the Electorate." November 8.

Page, Ann L., and Donald A. Clelland. 1978. "The Kanwha County Textbook Controversy: A Study of the Politics of Lifestyle Concern." *Social Forces* 59:169–85.

Plissner, Martin. 1983. "The Marriage Gap." *Public Opinion* (February-March): 53.

Plutzer, Eric. 1988. "Work Life, Family Life, and Women's Support of Feminism." *American Sociological Review* 53:640–49.

Plutzer, Eric. 1991. "Preferences in Family Politics: Women's Consciousness or Family Context?" *Political Geography Quarterly* 10. In press.

Rosenberg, Morris. 1968. *The Logic of Survey Analysis*. New York: Basic Books.

Weisberg, Herbert F. 1987. "The Demographics of a New Voting Gap: Marital Differences in American Voting." *Public Opinion Quarterly* 51:335–43.

Wildavsky, Aaron. 1987. "Choosing Preferences by Constructing Institutions: A Cultural Theory of Preference Formation." *American Political Science Review* 81(March): 3–22.

Wood, Michael, and Michael Hughes. 1984. "The Moral Basis of Moral Reform: Status Discontent vs. Culture as Explanations of Anti-Pornography Social Movement Adherence." *American Sociological Review* 49:86–99.

THE POLLS—POLL TRENDS

CHANGING AMERICAN ATTITUDES TOWARD THE SOVIET UNION

ALVIN RICHMAN

Changes in Soviet policies under Mikhail Gorbachev, U.S.-Soviet summitry, and the dismantling of the "iron curtain" in Eastern Europe have dramatically altered the way Americans view the Soviet Union. These changes are reflected in numerous survey measures, repeated during the past decade, that show that Americans' attitudes toward the USSR have changed from deep pessimism and hostility in the early 1980s, in the aftermath of the Soviet invasion of Afghanistan, to one of cautious optimism entering the 1990s. Most Americans, at this writing, see U.S.-Soviet relations greatly improved, and the USSR, under Gorbachev, no longer striving for "global domination." But at the same time, the USSR is still widely seen as a formidable rival for influence in various parts of the Third World. This mix of perceptions has resulted in the public's preference for a cautious, dual approach toward the USSR—showing receptivity to genuine Soviet conciliatory moves, but also keeping up our guard.

Public Perceptions of the Soviet Union

DIMINISHED PERCEPTION OF SOVIET THREAT TO THE UNITED STATES

Various polls show that the public's perceptions of the USSR as a military threat to the United States declined sharply in the late 1980s. For example, a survey by Americans Talk Security (ATS), taken in early 1990, found 33% of the public believed the Soviet Union posed either a "serious threat" (26%) or "very serious threat" (7%) to the United States—down from 60% in mid-1988. The predominant perception on ATS's latest survey is that the USSR represents a "minor threat" to the United States (49%). Only 16% said it was "not a threat

ALVIN RICHMAN is the Senior Public Opinion Analyst with the U.S. Department of State. The views expressed here are those of the author, not necessarily those of the Department of State.

Public Opinion Quarterly 55:135–148 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-01/$2.50

at all." Similarly, a separate ATS question asked about "how serious a threat" to U.S. security "during the next five years or so" was "Soviet aggression around the world." In early 1990, 24% responded either "extremely serious" (5%) or "very serious" (19%)—down markedly from the approximately 60% level that prevailed between October, 1987, and October, 1988. CBS/*New York Times* polls found that, by mid-1989, only 26% of the public believed the Soviet "military threat" posed a "real, immediate danger" to the United States—down from 54% in early 1985.

PERCEPTION OF SOVIET THREAT TO THE THIRD WORLD

During the late 1980s, Americans made a sharp distinction in their perceptions between the diminution of a direct Soviet threat to the United States and continued Soviet intervention in the Third World. A late 1987 ATS survey (not shown) found nearly two-thirds believed it was "likely" that the Soviet Union "will become more committed to reducing nuclear arms" (64%); however, a similar percentage thought it "unlikely" that the Soviets "will begin to pull back from their attempts to expand their influence in various parts of the Third World" (63%). In late 1989, an NBC/*Wall Street Journal* poll found that a majority believed it likely that the USSR would reduce its military presence in Eastern Europe (73%), and, also, that the "threat of war" would be reduced (78%). At the same time, a majority thought it unlikely that the USSR would end its "interference" in Central America (54%).

The widespread perception of a lessened direct Soviet threat to the United States appears to stem from reduced concern about both Soviet foreign policy objectives and Soviet military capabilities vis-à-vis the United States.

Soviet Objectives. During the past ten years, Roper polls have shown a steady decline in the percentage who perceive "global domination" as the USSR's main foreign-policy objective—from the peak of 73% in early 1980, to 58% in late 1984, shortly before Gorbachev's accession to power, and 25% in early 1990 (see fig. 1). Most of this decline has come in the group who hold the highly negative view that the Soviet Union would risk a "major war" to achieve domination (down from 39% in early 1980 to only 7% in 1990). In contrast, during this period, the percentage who viewed Soviet objectives as relatively limited increased steadily—from 23% in 1980 to 60% in 1990. Most of this increase has come in the group who see the Soviets' goal as "competition with the United States for more influence in different parts of the world" (up from 18% in 1980 to 44% in 1990).

Soviet Military Power. During the early 1980s, a plurality (about

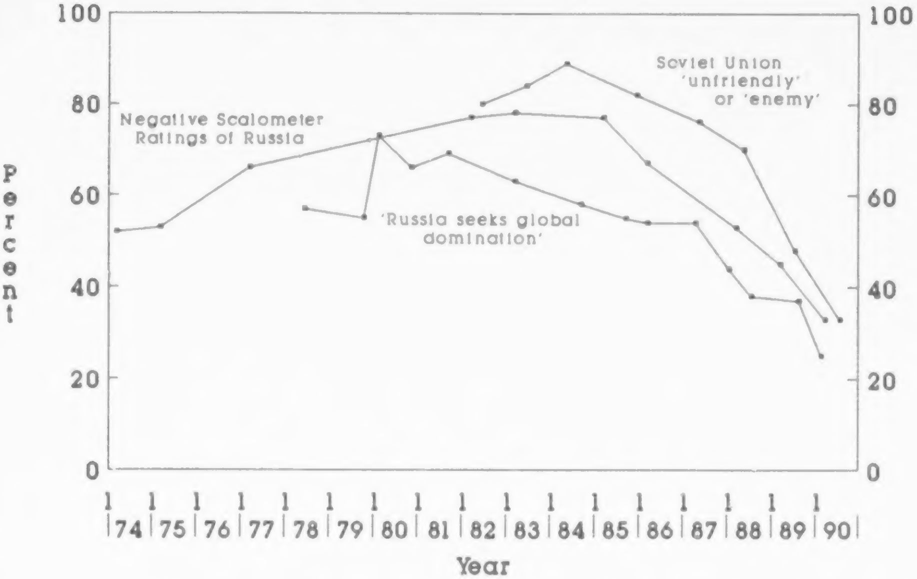


Figure 1. Changing Views of the Soviet Union

40% on the average) perceived the USSR militarily ahead of the United States, compared to about 15% who thought the United States was ahead and 35% who thought the two powers were "about equal." This has shifted in recent years. Now, a plurality of Americans (about 45% average) perceives the United States and USSR "about equal" in military strength, while roughly one-fourth think the USSR is ahead, and roughly one-fourth think the United States is ahead.

Despite these recent shifts, respect for Soviet military power is greater today than it was in the late 1940s, when a clear plurality viewed the United States ahead of the USSR (about 45% vs. about 20% who viewed the USSR as stronger). On the other hand, concern about Soviet "aggressive" intentions, which dropped sharply during the 1980s, is considerably less today than it was in the late 1940s.

U.S.-SOVIET RELATIONS SEEN AS THAWING

Ally-Enemy Ratings. Since 1982, Roper has regularly asked respondents to rate the Soviet Union (and other countries) on a five-point scale ranging from "close ally" to "enemy" of the United States. The ratings given the Soviet Union hit bottom in 1984, with 49% terming it an "enemy" of the United States and another 40% terming it "unfriendly." Negative views of Soviet policy have dropped sharply since then. In Roper's 1990 poll, 10% rated the Soviet Union as an "enemy" and 23% rated it "unfriendly" toward the United States (see fig. 1).

This decline among those viewing the USSR negatively is matched by a rise in the percentage viewing the USSR either as "neutral" toward the United States (up from 4% in 1984 to 31% in 1990) or as a "friend" of the United States (up from 2% to 20%).

The Cold War as "Ending," but Not Quite "Over." A number of pollsters have inquired about whether public perceptions of the cold war between East and West are keeping pace with actual events. Most of these polls show that optimism about a thaw in the cold war has risen during the past several years. An ABC/*Washington Post* poll in May, 1990, found 66% believed the cold war is "ending"—up from 50% one year earlier. However, the public is still closely divided about whether the cold war is "over," according to an October, 1990, poll by CBS/*New York Times* (49% viewed it as "over" vs. 43% said "not over").

GENERAL OPINION OF THE USSR

Perceptions of a lessened Soviet threat and improved relations with the USSR are reflected in an improvement in the overall image of the Soviet Union. Gallup and ABC/*Washington Post* polls in early 1990 found about three-fifths of Americans expressed a favorable opinion of the USSR—up from about one-fifth in the early 1980s. The current image of the USSR, in fact, is the most positive that has been recorded during the past four decades (see fig. 1).

Preferred Policy Toward the Soviet Union

While most Americans perceive U.S.-Soviet tensions as diminishing, the USSR is still predominantly viewed as a militarily powerful, global rival of the United States. This mix of perceptions has resulted in the public's preference for a two-way approach of being receptive to genuine Soviet offers while maintaining our guard.

PREFERENCE FOR A DUAL APPROACH

In recent years, Roper polls show little change in the public's predominant preference for a two-tiered approach of trying to improve relations with the USSR through negotiations while taking a "strong position" against Soviet expansion. About half of the public continues to favor this dual strategy (49%) according to Roper's February, 1990, survey.

However, the minority who favor a basically accommodating approach toward the USSR that stresses negotiating and avoidance of confrontation has risen somewhat since 1983 (from 23% to 29%). Conversely, the minority who favor a purely confrontational stance that stresses reliance on "increased military strength" to counter Soviet expansion has declined since 1983 (from 23% to 13%).

The widespread desire for caution in dealing with the USSR was also revealed by a November, 1989, *Time*-CNN/Yankelovich poll (not shown in the tables). A large majority (78%) preferred that President Bush "remain cautious in his dealings with the Soviets," rather than "act quickly to take advantage of recent changes in the Soviet Union" (16%).

PREFERRED AGREEMENTS WITH THE USSR

Various polls in recent years have shown that Americans give higher priority to reaching agreements with the Soviet Union on strategic nuclear arms control and cooperating on certain global issues (e.g., drugs and terrorism) than to reaching accords on other issues.

An August, 1989, Roper poll tested support for different arms control measures and found that the majority who favored a U.S.-Soviet agreement to "reduce long-range nuclear weapons by half" (77%) was larger than the majority who favored an agreement to "eliminate" U.S. and Soviet "short-range nuclear forces in Europe and Asia" (65%), or to "reduce" U.S. and Soviet conventional forces in Europe (60%).

Cooperating with the USSR on trade and other economic matters has lower priority than cooperating on arms control, regional conflicts, and prominent global issues, according to Roper's February, 1990, poll. Moreover, Americans clearly distinguish among the kinds of economic links with the USSR they favor and oppose. Over three-fifths *favored* selling grain and giving the USSR "favored nation trading terms"; three-fifths *opposed* providing loans or "foreign aid" to the USSR.

I am indebted to the following organizations for their research results; all samples are national representative telephone samples of the adult population unless otherwise noted.

SOURCES

ABC/*WP*: ABC/*Washington Post*
ATS: Americans Talk Security[1]

---

[1] Americans Talk Security national surveys were conducted alternately by one of three survey organizations: Market Opinion Research (later replaced by Market Strategies, Inc.), Daniel Yankelovich Group, and Marttila and Kiley, Inc. (see *Americans Talk Security Compendium*, comp. Marttila and Kiley, Inc. [Boston: Marttila and Kiley], 1989).

CBS/*NYT:* CBS/*New York Times*
Gallup: Gallup Organization (personal interviews used for scalometer ratings)
LAT: Los Angeles Times
MOR: Market Opinion Research
MS: Market Strategies, Inc.
M&K: Marttila and Kiley, Inc.
NBC/*WSJ:* NBC/*Wall Street Journal*
NORC/GSS: National Opinion Research Center, General Social Survey (personal interviews)
Roper: Roper Organization (personal interviews)
SRC: Survey Research Center, University of Michigan (personal interviews)
YANK: Daniel Yankelovich Group

Perceptions of the Soviet Union

DIMINISHED PERCEPTION OF SOVIET THREAT TO U.S.

ATS: How much of a threat would you say the Soviet Union is to the United States these days—a very serious threat, a serious threat, a minor threat, or not a threat at all?

	MS	*MOR*	*YANK*	*MOR*	*MOR*
	2–3/90	*12/88*	*6/88*	*4/88*	*1/88*
Very serious	7%	9%	18%	17%	16%
Serious	26	35	42	36	41
(Subtotal)	(33)	(44)	(60)	(53)	(57)
Minor	49	44	29	32	32
Not a threat	16	10	8	12	8
Don't know	2	2	3	3	3
N	1,000	1,000	1,006	1,000	1,000

ABC/WP: Same wording as ATS surveys above.

	6/89	*5/88*	*12/87*	*5/86*	*4/86*	*3/86*
Very serious	9%	11%	15%	9%	13%	20%
Serious	31	37	45	37	40	47
(Subtotal)	(40)	(48)	(60)	(46)	(53)	(67)
Minor	45	39	31	40	33	25
Not a threat	14	12	8	12	11	6
Don't know	1	2	1	2	2	2
N	1,500	1,508	1,007	NA	1,505	1,148

NOTE: NA = not available.

ATS: I'm going to read a list of several foreign policy issues and developments. Thinking about the next five years or so, please tell me whether you feel each issue I mention poses an extremely serious, very serious, somewhat serious, or not very serious threat to our country's national security interests. . . . Soviet aggression around the world.

	MS *2–3/90*	*YANK* *10/88*	*MOR* *9/88*	*M&K* *8/88*	*MOR* *4/88*	*YANK* *2/88*	*M&K* *10/87*
Extremely serious	5%	32%	21%	17%	22%	18%	21%
Very serious	19	37	31	33	40	42	35
(Subtotal)	(24)	(69)	(52)	(50)	(62)	(60)	(56)
Somewhat serious	39	21	31	33	28	30	28
Not very serious	35	8	14	15	8	8	13
Don't know	2	2	2	2	1	2	3
N	1,000	1,008	1,005	1,005	1,000	1,004	1,002

CBS/NYT: Do you believe the military threat from the Soviet Union is constantly growing and presents a real, immediate danger to the United States, or not?

	5/89	*2/85*	*9/83*	*4/83*
Yes	26%	54%	64%	57%
No	65	42	28	37
Don't know	9	5	8	7
N	1,073	1,533	1,587	1,489

PERCEPTION OF SOVIET THREAT TO THE THIRD WORLD

NBC/WSJ (12/89; *N* = 806 registered voters): Let me read you a number of things that people say might happen as a result of the recent events in Eastern Europe and the summit. [READ EACH ITEM AND ASK:] Do you think this is very likely to happen, somewhat likely to happen, or *not* that likely to happen?

	Very Likely	*Somewhat Likely*	*(Subtotal Likely)*	*Not Likely*	*Don't Know*
The United States and the Soviet Union will reduce the number of troops in Europe.	26%	47%	(73)	22%	5%
The threat of war will be reduced.	39	39	(78)	18	4
Soviet interference in Central America will end.	9	26	(35)	54	11

SOVIET OBJECTIVES

Roper: In your opinion, which of the following best describes Russia's primary objective in world affairs? (All *N*'s about 2,000.)

A. Russia seeks only to *protect* itself against the possibility of attack by other countries.
B. Russia seeks to *compete* with the U.S. for more influence in different parts of the world.
C. Russia seeks *global domination,* but not at the expense of starting a major war.
D. Russia seeks *global domination* and will *risk a major war* to achieve that domination if it can't be achieved by other means.

	Limited Objectives			*Global Domination*[a]			
	Protect	*Compete*	*(Subtotal)*	*Low Risk*	*High Risk*	*(Subtotal)*	*Don't Know*
2/90	16%	44%	(60)	18%	7%	(25)	14%
8/89	12	41	(53)	26	11	(37)	10
7/88	10	40	(50)	29	9	(38)	11
1/88	12	34	(46)	33	11	(44)	9
4–5/87	9	31	(40)	37	17	(54)	5
3/86	8	30	(38)	36	18	(54)	9
9/85	8	29	(37)	35	20	(55)	9
9/84	7	28	(35)	38	20	(58)	8
3/83	9	23	(32)	34	29	(63)	5
9/81	6	18	(24)	35	34	(69)	7
11/80	5	21	(26)	38	28	(66)	8
2/80	5	18	(23)	34	39	(73)	4
10–11/79	9	30	(39)	37	18	(55)	6
6/78	7	25	(32)	34	23	(57)	11

[a] See figure 1.

CBS/NYT and Roper/Fortune: Generally, would you describe the Soviet Union as a peace-loving nation willing to fight only if it (she) thinks it has to defend itself, or as an aggressive nation that would start a war to get something it (she) wants? (Roper polls—6/47 and 6/45—used the term *she*.)

	10/90	*11/89*	*5/89*	*9/85*	*6/47*	*6/45*
Aggressive	38%	42%	57%	69%	68%	38%
Aggressive, but not war[a]	8	12	7	5	6	8
(Subtotal)	(46)	(54)	(64)	(74)	(74)	(46)
Peace loving	43	36	27	17	13	39
Don't know	11	11	10	9	13	15
N	960	1,297	1,073	1,277	NA	NA

[a] The response "Aggressive, but not war" was volunteered.

SOVIET MILITARY POWER

CBS/NYT: Right now, would you say the United States is superior in military strength to the Soviet Union, or is about equal in strength, or is not as strong as the Soviet Union?

	U.S. Not As Strong	*About Equal*	*U.S. Superior*	*Don't Know*	*N*
5/90	26%	44%	21%	9%	1,140
11/89	27	47	18	8	1,297
5/89	28	50	17	5	1,073
5/88	29	45	14	12	939
9/87	31	44	14	11	836
11/85	25	47	18	10	800
1/85	29	46	17	8	1,534
6/83	31	44	13	13	1,365
4/83	42	36	12	10	1,489
1/83	32	47	11	11	1,597
1/82	44	37	9	9	1,540
6/81	42	39	11	8	1,433

	U.S. Not As Strong	*About Equal*	*U.S. Superior*	*Don't Know*	*N*
1/80	42	34	14	11	1,468
6/79	43	30	11	16	1,422
6/78	42	32	12	14	1,527

ABC/WP: As things presently stand, who do you think is ahead in military power, the United States or the Soviet Union, or do you think they are about the same in military strength?

	USSR Ahead	*About Equal*	*U.S. Ahead*	*Don't Know*	*N*
5/90	21%	52%	26%	1%	1,526
11–12/87	31	47	20	2	1,007
11/85	26	46	24	4	1,507
1/85	22	55	18	5	504
11/83	29	47	17	6	1,505
9/83	37	42	17	5	1,506
7–8/83	38	45	13	4	1,505
10/81	40	34	19	6	1,505
4/81	41	36	18	6	1,515
6/79	31	45	19	6	839

SRC: How would you say that Russia compares with the United States in military power—is Russia stronger than the United States, just as strong, or weaker?

	Russia Stronger	*Equal*	*Russia Weaker*	*Don't Know*	*N*
11–12/49	19%	21%	40%	20%	606
10/48	20	12	47	21	605
12/46–1/47	16	12	47	25	571

U.S.-SOVIET RELATIONS SEEN AS THAWING

Roper: "I'd like to have your impressions about the overall position that some countries have taken toward the U.S. [Respondents shown list of countries.] Would you read down that list, and for each country, tell me if you believe the country has acted as a close ally of the U.S., has acted as a friend but not close ally, has been more or less neutral toward the U.S., has been mainly unfriendly toward the U.S., but not an enemy, or has acted as an enemy of the U.S.? . . . The Soviet Union. (All *N*s about 2,000.)

	Close Ally	*Friend*	*(Positive Subtotal)*	*Neutral*	*Unfriendly*	*Enemy*	*(Negative Subtotal)*[a]	*Don't Know*
7/90	2%	20%	(22)	31%	23%	10%	(33)	13%
7/89	1	15	(16)	27	34	14	(48)	9
5/88	1	6	(7)	16	40	30	(70)	8
5/87	—	4	(4)	13	44	32	(76)	7
12/85	1	2	(3)	9	42	40	(82)	6
5–6/84	1	2	(3)	4	40	49	(89)	4
6/83	1	2	(3)	8	43	41	(84)	5
6/82	1	3	(4)	10	40	40	(80)	6

[a] See figure 1.

ABC/WP: Would you agree or disagree that the cold war between East and West is coming to an end?

	Agree	*Disagree*	*Don't Know*	*N*
5/90	66%	32%	3%	1,526
3–4/90	55	31	14	1,000
11/89	58	30	12	1,007
6/89	47	48	5	1,500
5/89	50	46	5	1,513

LAT: Do you think the cold war is over, or not?

	Yes	*No*	*Don't Know*	*N*
5/90	36%	51%	13%	2,144
11/89	22	61	17	1,618
7/88	15	70	15	2,277

CBS/NYT: Ever since the 1940s, people have called relations between the United States and the Soviet Union "the cold war." In view of recent events in the Soviet Union and Eastern Europe, do you think the cold war is *over*?

	Yes	*No*	*Don't Know*	*N*
10/90	49%	43%	8%	960
6/90	41	52	7	1,084
3–4/90	34	56	10	1,515
1/90	37	54	9	1,557
11/89	37	54	9	1,297

GENERAL OPINION OF THE USSR

Gallup: I'd like your opinion of some foreign countries. Is your overall opinion of the Soviet Union very favorable, mostly favorable, mostly unfavorable, or very unfavorable?

	Favorable			*Unfavorable*				
	Very	*Mostly*	*(Positive Subtotal)*	*Mostly*	*Very*	*(Negative Subtotal)*	*Don't Know*	*N*
9/90	7%	51%	(58)	23%	9%	(32)	10%	1,031
5/90	7	48	(55)	24	8	(32)	13	1,255
2/90	13	51	(64)	18	8	(26)	10	1,235
8/89	8	43	(51)	27	13	(40)	9	1,250
7/89	6	40	(46)	29	14	(43)	11	1,000
2–3/89	7	55	(62)	20	9	(29)	9	1,000

ABC/WP: Do you generally have a favorable or unfavorable opinion of the following countries? The Soviet Union.

	Favorable	*Unfavorable*	*Don't Know*	*N*
3/90	60%	38%	2%	760
3–4/89	55	44	1	1,525
2/89	45	52	3	1,512

NORC/GSS—Scalometer Ratings of Russia: You will notice that the boxes on this card go from the highest position of plus five for a country which you like very much to the lowest position of minus five for a country which you dislike very much. How far up the scale or how far down the scale would you rate Russia? (Don't know responses are not included.)

		Favorable			*Unfavorable*		
	Very (+5, +4)	*Somewhat (+3, +2, +1)*	*(Positive Subtotal)*	*Somewhat (−1, −2, −3)*	*Very (−4, −5)*	*(Negative Subtotal)*[a]	*N*
2–4/90	11%	55%	(66)	19%	15%	(34)	867
2–4/89	9	46	(55)	24	21	(45)	951
2–4/88	5	42	(47)	26	27	(53)	930
2–4/86	2	31	(33)	33	34	(67)	1,405
2–4/85	3	20	(23)	32	45	(77)	1,450
2–4/83	2	21	(23)	34	44	(78)	1,528
2–4/82	2	20	(22)	26	51	(77)	1,445
2–4/77	4	30	(34)	28	38	(66)	1,415
2–4/75	6	41	(47)	26	27	(53)	1,386
2–4/74	6	42	(48)	25	27	(52)	1,411

[a] See figure 1.

Gallup—Scalometer Ratings of (Russia) (The Soviet Union): Notice that the boxes on this card go from the highest position of +5 for a country you have a very favorable opinion of all the way down to the lowest position of −5 for a country you have a very unfavorable opinion of. How far up the scale or how far down the scale would you rate (Russia) (The Soviet Union)? (Personal interviews.)

		Favorable			*Unfavorable*			
	Very (+5, +4)	*Somewhat (+3, +2, +1)*	*(Positive Subtotal)*	*Somewhat (−1, −2, −3)*	*Very (−4, −5)*	*(Negative Subtotal)*	*Don't Know*	*N*
Soviet Union								
4–5/87	2%	23%	(25)	38%	33%	(71)	4%	4,244
8/82	4	17	(21)	27	47	(74)	5	1,500
1/80	2	11	(13)	22	62	(84)	3	1,597
Russia								
6/76	4	17	(21)	26	46	(72)	7	1,544
7/73	9	36	(45)	23	21	(44)	11	1,544
5/72	6	34	(40)	22	32	(54)	6	1,541
12/66	2	15	(17)	28	48	(76)	7	3,542
12/56	1	4	(5)	15	72	(87)	8	1,588
8/54	1	4	(5)	12	76	(88)	7	1,593
9/53	—	1	(1)	16	76	(92)	7	1,488

Preferred Policy Toward the USSR

PREFERENCE FOR A DUAL APPROACH

Roper: In light of Russia's recent actions, which of these statements best describes your opinion about the future direction of U.S.-Russian relations? (All *N*s about 2,000.)

A. *Confrontational Approach:* It's clear that Russia can't be trusted and we will have to rely on increased military strength to counter them in the future.

B. *Dual Approach:* We should take a strong position with the Russians now so they won't go any further, but at the same time we should try to reestablish good relations with them.

C. *Accommodating Approach:* We should do nothing that is likely to provoke a U.S.-Russian military conflict, but instead try to negotiate and reason out our differences.

	Accommodating Approach	*Dual Approach*	*Confrontational Approach*	*Don't Know*
2/90	29%	49%	13%	10%
8/89	22	58	13	7
7/88	27	53	13	8
1/88	23	55	14	8
4–5/87	18	57	20	4
3/86	23	54	16	7
9/85	21	54	19	6
9/84	22	54	18	6
3/83	23	50	23	5

PREFERRED AGREEMENTS WITH THE USSR

Roper: There are a number of specific issues on which the United States and the Soviet Union might try to work out an agreement. Here is a list of some possible agreements that could be negotiated. [Card shown to respondent.] Would you please read down that list and for each one tell me if you would favor or oppose such an agreement? (Don't know responses are omitted.)

	August, 1989		*January, 1988*		*April-May, 1987*	
	Favor	*Oppose*	*Favor*	*Oppose*	*Favor*	*Oppose*
An agreement to reduce U.S. and Russian long-range nuclear weapons by half.	77%	14%	72%	17%	72%	20%
An agreement that bans the testing of nuclear weapons by both the U.S. and the Soviet Union.	67	23	67	21	67	23
An agreement to eliminate U.S. and Soviet short-range nuclear forces in Europe and Asia.	65	22	NOT ASKED		NOT ASKED	
An agreement to eliminate U.S. and Soviet medium-range nuclear weapons in Europe and Asia.	NOT ASKED		64	23	62	28
An agreement to have with the Soviet Union the same favored trade terms we						

	August, 1989		*January, 1988*		*April-May, 1987*	
	Favor	*Oppose*	*Favor*	*Oppose*	*Favor*	*Oppose*
have with certain other nations, like Britain, France, Canada, etc.	65	20	59	23	58	29
An agreement to reduce the number of U.S. and Soviet non-nuclear military forces stationed in Europe.	60	28	NOT ASKED		NOT ASKED	
An agreement to reduce the number of U.S. and soviet troops stationed in central Europe.	NOT ASKED		56	29	57	31
N	1,991		1,982		2,000	

Roper (2/90; *N* = 1,977): Here is a list of a number of ways we could cooperate with the Soviet Union. Would you please read down the list and for each item tell me if it is one of the most important things we should do with the Soviet Union, or if it would be important to do but not a top priority, or something we shouldn't do with them.

	Most Important	*Not Top Priority*	*(Subtotal)*	*Should Not Do*	*Don't Know*
Work together to fight terrorism around the world.	75%	18%	(93)	4%	3%
Work together to improve the environment.	74	20	(94)	3	3
Join together in a peace-keeping force to prevent conflicts around the world.	65	23	(88)	9	3
Work together to resolve the Arab-Israeli conflict in the Middle East.	39	41	(80)	12	9
Sharply reduce the number of U.S. and Soviet troops in Europe.	38	35	(73)	20	8
Sign a long-term agreement to trade our grain for their oil and natural gas.	32	37	(69)	22	9
Give the Soviet Union the same favored nation trading terms we have with countries					

	Most Important	*Not Top Priority*	*(Subtotal)*	*Should Not Do*	*Don't Know*
like Canada, the Western European nations, and Japan.	26	40	(66)	21	13
Provide the Soviet Union with technical and business assistance to help rebuild their industry.	19	42	(61)	32	8
Encourage American companies to invest in the Soviet Union.	16	37	(53)	39	9
Have the U.S. government provide foreign aid to the Soviet Union.	7	26	(33)	59	7
Encourage American banks to provide loans to the Soviet Union.	7	25	(32)	61	8

Public Opinion Quarterly

Published for the American
Association for Public Opinion
Research by the University of
Chicago Press

ELEANOR SINGER, THERESA F. ROGERS, and MARC B. GLASSMAN	Assessing an Information Campaign on AIDS
KIM FRIDKIN KAHN and EDIE N. GOLDENBERG	Women Candidates in the News
PETER S. TUCKEL and BARRY M. FEINBERG	Answering Machines and Telephone Surveys
JOHN G. GEER	The Electorate's Partisan Evaluations
GARY A. MAUSER and CRAIG FITZSIMMONS	The Effect of Election Polls on Exchange Rates
RICHARD D. DIXON and others	Self-Interest and Public Opinion on Smoking
ROGER TOURANGEAU, KENNETH A. RASINSKI, and NORMAN BRADBURN	Measuring Happiness in Surveys
GABRIEL WEIMANN	Back to the Concept of Opinion Leaders?
IN MEMORIAM	I. A. Lewis
THE POLLS	Review: Which Side Are You On? *by Peter V. Miller*
BOOKS	Reviews and Book Notes *by Alex Inkeles, Andrew Kohut, and Daisy Quarm*

Summer 1991

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR Advisory Committee

Eleanor Singer, Chair (1987–91)
Herbert Abelson (1988–92) Gladys Engel Lang (1989–93)
W. Russell Neuman (1990–94)

Stanley Presser and Kathleen Frankovic,
AAPOR Councillors-at-Large

Howard Schuman, *Editor*

Manuscript Editor John D. Vander Weg
Office Manager Carlotta Nosel
Assistant to the Editor Katherine Bischoping

The Polls
Reviews: Stanley Presser, Seymour Sudman
Trends: Robert Y. Shapiro

EDITORIAL BOARD

Herbert I. Abelson
Robert P. Abelson
Sandra J. Ball-Rokeach
Larry M. Bartels
Nathaniel Beck
Sandra H. Berry
Lawrence D. Bobo
Pamela Conover
Jean M. Converse
Stanley Feldman
Richard J. Fox
Deborah R. Hensler
Shanto Iyengar
Graham Kalton
Gary King
Jon A. Krosnick
Richard A. Kulka
Kurt Lang
McKee J. McClendon
William J. McGuire
Michael MacKuen
Jack M. McLeod
Elizabeth Martin
John Mueller
Harold Nisselson
Helmut Norpoth
Vincent Price
Willard L. Rodgers
Norbert Schwarz
Robert Y. Shapiro
Lee Sigelman
Tom W. Smith
John Sullivan
John Zaller

Public Opinion Quarterly (ISSN 0033-362X) Volume #55, 4 issues (1991)

© 1991 by the American Association for Public Opinion Research.

Published quarterly: Spring, Summer, Autumn, and Winter, by The University of Chicago Press, 5720 S. Woodlawn Avenue, Chicago, Illinois 60637. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $40.00; individuals 1 year $20.00. WAPOR member subscription rate, U.S.A.: individuals 1 year $18.00. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $18.00. Other countries add $3.00 for each year's subscription to cover postage.

Japanese subscription agent: Kinokuniya Company Ltd.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to: Journals Fulfillment Department, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Institute for Social Research
P.O. Box 1248, Ann Arbor, MI 48106
Business, Subscriptions, Back Issues: The University of Chicago Press,
P.O. Box 37005, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, New Jersey 08540

Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. ♾

PUBLIC OPINION QUARTERLY

VOLUME 55 / SUMMER 1991 / NUMBER 2

Articles

Public Opinion about AIDS Before and After the 1988 U.S. Government Public Information Campaign 161
ELEANOR SINGER, THERESA F. ROGERS, and MARC B. GLASSMAN

Women Candidates in the News: An Examination of Gender Differences in U.S. Senate Campaign Coverage 180
KIM FRIDKIN KAHN and EDIE N. GOLDENBERG

The Answering Machine Poses Many Questions for Telephone Survey Researchers 200
PETER S. TUCKEL and BARRY M. FEINBERG

The Electorate's Partisan Evaluations: Evidence of a Continuing Democratic Edge 218
JOHN G. GEER

The Short-Term Effect of Election Polls on Foreign Exchange Rates: The 1988 Canadian Federal Election 232
GARY A. MAUSER and CRAIG FITZSIMMONS

Self-Interest and Public Opinion toward Smoking Policies: A Replication and Extension 241
RICHARD D. DIXON, ROGER C. LOWERY, DIANE E. LEVY, and KENNETH F. FERRARO

Measuring Happiness in Surveys: A Test of the Subtraction Hypothesis 255
ROGER TOURANGEAU, KENNETH A. RASINSKI, and NORMAN BRADBURN

The Influentials: Back to the Concept of Opinion Leaders? 267
GABRIEL WEIMANN

In Memoriam

I. A. Lewis, 1921–1990 280

The Polls

Review: Which Side Are You On? The 1990 Nicaraguan Poll Debacle 281
PETER V. MILLER

Book Reviews

Ronald Inglehart, *Culture Shift in Advanced Industrial Society* 303
ALEX INKELES

Paul R. Abramson, John H. Aldrich, and David Rohde, *Change and Continuity in the 1988 Elections* 307
ANDREW KOHUT

Murray A. Straus and Richard J. Gelles, *Physical Violence in American Families: Risk Factors and Adaptions to Violence in 8,145 Families* 308
DAISY QUARM

Book Notes 311

NOTICE TO CONTRIBUTORS

Prospective contributors should refer to the detailed Notice to Contributors that appears in the Spring 1991 issue of *POQ*

PUBLIC OPINION ABOUT AIDS BEFORE AND AFTER THE 1988 U.S. GOVERNMENT PUBLIC INFORMATION CAMPAIGN

ELEANOR SINGER,
THERESA F. ROGERS, and
MARC B. GLASSMAN

Abstract In October, 1987, the Centers for Disease Control mounted a massive public information campaign to alert the public to the dangers of AIDS and to provide information about its transmission and prevention. Using data from two Gallup surveys, one just before the campaign began and the other several months after its conclusion, we examine changes in public information and misinformation about transmission, in concern about AIDS as an epidemic, and in reported behavior to avoid exposure to AIDS. We conclude that although some changes in knowledge did take place, these were essentially a continuation of trends beginning before the public information campaign and continuing well after its conclusion. For these and other reasons, we argue that the effects of the campaign on public information were minimal. However, between 1987 and 1988 there was a small but statistically significant increase in reported condom use, an increase paralleled by increased condom sales between 1986 and 1988. In addition, there was a substantial increase in the number of people expressing concern about AIDS as an epidemic for the population at large. The campaign may well have contributed to both of these changes.

Beginning in October, 1987, the Centers for Disease Control (CDC) mounted a massive campaign to alert the public to the dangers of

ELEANOR SINGER and THERESA F. ROGERS are Senior Research Scientists at Columbia University; MARC B. GLASSMAN is a Senior Statistician at the Port Authority of New York and New Jersey. The authors are grateful for the support of the American Foundation for AIDS Research (Grant #000382), which made possible the analyses reported here. We would like to thank George Shambaugh for his help in processing the data. The detailed and helpful editorial suggestions of Diana Mutz and Vincent Price are also gratefully acknowledged.

Public Opinion Quarterly 55:161–179 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-02/$2.50

human immunodeficiency virus (HIV), the cause of acquired immunodeficiency syndrome (AIDS), and to provide information about its transmission and the prevention of AIDS. The centerpiece of this multimillion dollar campaign, designed by the advertising agency Ogilvy and Mather, was the eight-page brochure, "Understanding AIDS," from Surgeon General C. Everett Koop; the brochure was mailed to every household in the country in late May and early June of 1988. The campaign also included radio and television spots and public service announcements aimed at specific subgroups (teenagers, sexually active adults, women at high risk of infection). The pamphlet emphasized the behaviors that put one at risk of contracting AIDS. Although it also provided information about how AIDS *cannot* be transmitted, the major focus was on infection and how to protect against it. In this article, we appraise the effects of the government information campaign on public knowledge, attitudes, and behavior in relation to AIDS.

We cannot, in this brief article, do justice to the extensive literature on public information campaigns. Hyman and Sheatsley's early analysis (1947), "Why Information Campaigns Fail," has been transformed, by Mendelsohn (1973) and subsequent researchers, into prescriptions for their success. But that success is known to be limited. In general, it is believed to be easier to communicate information than to change attitudes, and easier to change attitudes than to bring about changes in behavior that is deeply ingrained or that serves important psychological functions (Nelkin 1987; Rogers and Storey 1987). Indeed, changes in knowledge and attitudes, and an absence of change in behavior, were documented in the evaluation of a year-long public information campaign about AIDS in Great Britain, initiated about two years before the one in the United States (Department of Health and Social Security and the Welsh Office, 1987).[1] But as we show below, results from the U.S. campaign were more equivocal.

Data and Methods

Our analysis of the changes that took place is based on two Gallup Poll personal interview surveys, one done in October, 1987, and the other in October, 1988.[2] The 1987 questionnaire contained nine questions about AIDS, many with subquestions. Intended for dissemination to the news media, these questions inquired into knowledge about transmission, concern about spread, behavior changes, attitudes toward people with AIDS, and attitudes toward AIDS-related policies.

---

[1] To be precise, the evaluation found claimed changes in behavior, but no change when responses to behavior questions were compared.

[2] The Gallup Poll is an omnibus survey containing questions on different topics. It is based on a probability sample to the block level; at that point, interviewers complete specified quotas after making a random start.

We commissioned some of these questions to be asked again in October, 1988. Because we could not afford to replicate them all, we selected those we felt would be most useful in appraising change as a result of the government information campaign.[3] We decided to look at changes in knowledge (measured by eight questions concerning routes of HIV transmission), concern about AIDS (measured by two questions, one asking about personal worry, the other about concern that AIDS would become an epidemic for the population at large), and reported behavior change (measured by three questions about condom use, avoiding homosexuals, and avoiding medical treatments requiring blood transfusions). We also included one question about whether employers should be free to fire people with AIDS, which had been asked by Gallup twice before. The exact wording of all 14 questions, which constitute the dependent variables of the analysis, appears in the appendix.

In order to assess change in these variables over time, we cross-classified (i.e., "predicted") each of them by study (pre- and postcampaign Gallup surveys), to determine whether there were any overall campaign effects, and by a set of demographic variables (age, sex, education, marital status, and race), to determine whether differential effects obtained for population subgroups. Then, to assess the significance of these changes, we fit logit models to the cross-tabulations. Such models are the preferred form of log-linear analysis when, as here, one of the variables is clearly dependent and the others are conceived as predictor variables. We used log-linear analysis because all the variables of interest except age are categorical (Bishop, Fienberg, and Holland 1975; Goodman 1978).

The logit models include study (which is a proxy for time), age, education, and either race, gender, or marital status as independent variables. Because of the small number of cases, we could not include both race and marital status in the same models, and, therefore, we ran two and sometimes three sets of models: one including marital status or gender along with age, education, and study, the other including race and the same three variables. Furthermore, again because of the small number of cases, we collapsed education and age into two categories: high school graduation or less versus more than high school; and 18–44 versus 45 and older.

---

[3] A full-scale evaluation of the campaign would have required many more resources than we had. In particular, it would have been desirable to know whether respondents remembered having seen the Surgeon General's pamphlet or seen or heard any of the campaign's public service announcements, but we had no money for the detailed questioning that would have been required. As a result, we can say something about the changes that occurred, but cannot link them unequivocally to exposure to the campaign itself.

The analysis involved two steps. First, we found the best-fitting logit model for each of the fourteen dependent variables and for each combination of independent variables—30 models in all.[4] Then, in order to test the significance of changes over time in the sample as a whole and in subgroup responses, we deleted, one at a time, the main effect for study and the relevant first-order interaction terms from the best-fitting model and evaluated the significance of the resulting change in chi-square.[5] For example, the best-fitting model for responses to the question about condom use included all main effects, all three-way effects, and the interaction between condom use, race, education, and study; we deleted the main effect for study (CS) and the interactions (three-way effects) between condom, education, and study (CES), condom, age, and study (CAS), and condom, race, and study (CRS) one at a time from this model and found that none of these terms was significant (see table 2, D12, Model RASEC).

Changes in the Sample as a Whole

We begin by examining changes in public opinion about AIDS in the sample as a whole and then turn to changes in demographically defined subgroups. Table 1 presents the raw data for the 1987 and 1988 surveys.

Column 1 shows the proportion of each sample giving the specified response to each of the fourteen questions. Subgroup responses are shown in columns 2 through 6: age in column 2; race in column 3; education in column 4; marital status in column 5, and gender in column 6. The first row shown for each question displays 1987 data; the second row, 1988 data.

So far as actual causes of AIDS are concerned, we found that knowledge that heterosexual contact could transmit the HIV virus increased by 4.5 percentage points, from 87.7% in 1987 to 92.2% in 1988. Knowledge about homosexual transmission and transmission by sharing needles, which were already close to 100%, remained virtually unchanged during this period.

---

[4] Because of the large number of models involved, we adopted a uniform screening procedure to help simplify the search for the best-fitting model. We began by looking at the independence model, a model including all main effects, a model including all three-way effects, and a model including all four-way effects. Once we found a level that produced an acceptable fit—i.e., a model that fit the data at a probability of at least .20—we worked backward, deleting any variable at that level that did not have a statistically significant parameter associated with it.

[5] By "relevant" we mean first-order interactions between each of the demographic characteristics and study, which is a proxy for time. When we could not test the significance of an interaction term in this way, because the term did not appear in the best-fitting model, we used, as a guide to statistical significance, the parameter estimates associated with that particular term in the model including all first-order interactions. We did not test the significance of interactions between demographic characteristics.

So far as "noncauses" of AIDS are concerned, we found that the percentage believing the HIV virus could be transmitted by sneezing (plus the percentage unsure of the correct response) dropped by 6.7 percentage points;[6] by a drinking glass, 6 percentage points; by working in the same room, 5.6 percentage points. Thus, looking at information gain in the sample as a whole, we found sizable change on four of the eight knowledge questions asked. Although only 10.7% of the sample believed, by October, 1988, that AIDS could be transmitted by someone working in the same room, misinformation concerning the other causes remained substantial, at around 30%.

For the sample as a whole there were no changes in the course of the year in the belief that employers should have the right to fire workers who have AIDS, nor in personal concern about AIDS. Concern that AIDS might become an epidemic for the public at large, on the other hand, increased more than any other item.[7]

Table 2 presents, in column 1, the models tested for each question (or dependent variable); in column 2, the best-fitting model; and in column 3, the terms found to be significant when deleted from the best-fitting model.

For the sample as a whole, we found significant change (that is, a significant main effect of study, or time) on 9 of 14 dependent variables (see table 2, col. 3): Five knowledge questions (heterosexual contact [Hetersex], sneezing [Sneezing], sharing a drinking glass [Drink], donating blood [Donate], and working next to someone [Work]); one

---

[6] In all of these comparisons, DK responses are grouped with the *wrong* response.

[7] There are three possible explanations for this large increase. One is that the information campaign itself increased public concern about the extent of the danger posed by AIDS. The other is that we inadvertently increased the proportion responding "yes" to this question by eliminating several groups that were asked about in 1987 in addition to "the population at large." Even though "the population at large" was the first such group asked about in 1987, encouraging us to believe that we could omit the others, it was possible that some of the other groups (homosexuals, drug users) provided a contrast effect in 1987, thus reducing the percentage of people saying they considered AIDS a threat to the population at large.

These alternatives were tested by Gallup in 1989 in a split-ballot experiment, in which half the sample was asked the original question and half was asked the version we had used in 1988. Differences between the two versions were not statistically significant, thus suggesting that the increase between 1987 and 1988 was substantive, and not artifactual. See Singer, Colasanto, and Rogers 1990.

Still a third possibility, however, is that the 1987 figure was artificially low for other reasons, since other polls taken around the same time all indicate that about 70% of the public was concerned that AIDS might spread to the general population.

Table 1. Public Opinion about AIDS, Gallup October, 1987, and October, 1988 (percent giving specified response)

			Age (2)		Race (3)		Education (4)		Marital Status (5)		Gender (6)	
Question[a]	Response	Total (1)	18–44	45 +	White	Nonwhite	HS or less	More than HS	Married	Not Married	Male	Female
						A. Knowledge about Transmission						
1. Homosex	Yes	94.8	94.4	95.5	94.9	94.4	94.9	94.8	95.4	93.9	—	—
		94.7	95.8	93.1	95.4*	89.5	93.0*	96.6	94.8	94.4	—	—
2. Hetersex	Yes	87.7	86.7	89.1	87.7	87.1	88.7	86.1	87.2	88.5	—	—
		92.2	93.5*	90.5	92.7	88.6	90.5*	94.1	92.8	91.0	—	—
3. Needles	Yes	97.2	97.6	96.5	97.2	97.2	97.7	96.5	97.6	96.5	—	—
		96.0	97.9*	93.4	96.6*	91.5	93.0	99.5	96.2	96.7	—	—
4. Toilet	Yes/DK	28.0	25.7*	31.4	26.7*	36.4	33.5*	19.3	26.3	31.0	—	—
		26.9	22.8	32.4	24.5*	43.2	34.6*	17.9	27.2	26.5	—	—
5. Drink	Yes/DK	34.1	30.5*	39.8	33.8	36.3	39.2*	26.2	33.6	35.1	—	—
		28.1	22.6*	35.1	25.9*	42.0	34.9*	20.0	28.8	26.8	—	—
6. Donate	Yes/DK	33.6	30.1*	39.0	32.1*	43.9	38.7*	25.5	33.0	34.6	—	—
		34.5	30.0	40.3	31.6*	53.3	44.6*	22.4	32.6	37.9	—	—
7. Sneezing	Yes/DK	34.0	30.8*	38.9	33.8	35.3	38.4*	27.0	31.7*	38.1	—	—
		27.3	22.4*	33.5	24.9*	42.7	30.5	23.4	27.0	27.8	—	—
8. Work	Yes/DK	16.3	14.5*	19.2	16.2	17.4	19.6*	11.3	16.6	15.8	—	—
		10.7	7.4*	15.0	8.8*	24.2	15.0*	5.5	11.5	9.3	—	—

						B. Concern about Spread[a]						
9. Epidemic	Likely	54.4	55.6	52.4	53.0*	64.4	57.6*	49.5	52.7*	57.6	—	—
		75.4	74.5	76.6	74.2*	84.2	79.4*	71.0	53.0	64.4	—	—
10. Concerned	Very	20.0	23.6*	14.6	17.8*	35.4	21.4	17.9	17.8	24.2	—	—
		18.9	22.0	14.8	15.5*	41.1	22.2	14.9	16.6	22.9	—	—
						C. Attitudes toward Regulation of Others						
11. Fire	Agree	24.7	22.3*	28.4	25.0	22.8	28.8*	18.4	24.5	25.1	—	—
		25.6	20.5*	32.2	27.1	15.3	28.2*	22.5	27.7*	21.8	—	—
						D. Reported Behavior Change						
12. Condom	Yes	8.4	12.9*	1.5	7.1*	17.5	7.5*	10.0	—	—	11.4*	5.5
		11.6	17.5*	3.8	10.9	16.0	8.2*	15.5	—	—	12.7	10.5
13. Avoidhom	Yes	7.8	9.4*	5.3	6.6	9.8	7.9	7.5	—	—	9.5*	4.3
		11.2	13.1*	8.6	11.3	12.6	11.8*	10.5	—	—	13.8*	8.9
14. Avoidmed	Yes	9.8	10.5	8.7	9.4	12.5	10.3	9.1	—	—	9.1	10.5
		13.8	13.7	13.8	12.6*	21.4	14.4	13.0	—	—	10.9*	17.1

NOTES: The first row of percentages for each item lists 1987 data, the second row lists 1988 data; $N = 1{,}569$ for the 10/87 survey and $N = 1{,}027$ for the 10/88 survey. Both surveys are based on national samples and were done in person. The percentages shown are derived from data weighted by Gallup to adjust for "times at home" and to align sample distributions with population parameters derived from the U.S. Census; all percentages are based on respondents with no missing data.

[a] See the appendix for exact question wording.

* $p < .05$, two-tailed chi-square for differences between categories of a demographic variable.

Table 2. Summary of Logit Analyses, Gallup October, 1987, and October, 1988

Dependent Variable[a]	Model Tested[b] (1)	Best-Fitting Model ($p \geq .20$) (2)	Significant Interactions with Time (S)[c] (3)
		A. Knowledge about Transmission	
1. Homosex (H)	GASEH	Independence model (I) + main effects (M) + 3-way interaction between homosex, age, and study (HAS)	HAS
	RASEH	Saturated model	—
	MASEH	I + M + HAS + HAM	HAS
2. Hetersex (H)	RASEH	Saturated model	HS, HAS
	MASEH	I + M + all 3-way interactions (3-ways) + HEMS	HS, HAS, HES
3. Needles (N)	RASEN	Saturated model	NES
	MASEN	I + M + NES	NES
4. Toilet (T)	RASET	I + M + TAE + TAR	—
	MASET	I + M + TAE + TMS	TMS
5. Drink (D)	RASED	I + M	DS
	MASED	I + M + DEM	DS
6. Donate (D)	RASED	I + M + 3-ways + DARS	DS, DAS, DES, DRS
	MASED	I + M + DES + DEM	DES
7. Sneezing (S)	RASES	Saturated model	SAS, SES, SRS
	MASES	I + M + 3-ways + SAMS	SS
8. Work (W)	RASEW	I + M + WRS	WRS
	MASEW	I + WS + WA + WE	WS
		B. Concern about Spread	
9. Epidemic (E)	GASEE	I + M + 3-ways + EAES	ES
	RASEE	I + M + 3-ways + EAES	ES
	MASEE	I + M + 3-ways + EAES	ES
10. Concerned (C)	RASEC	Saturated model	CAS
	MASEC	I + M + CAM + CAS	CAS, CES
		C. Attitudes toward Regulation	
11. Fire (F)	RASEF	Saturated model	FAS
	MASEF	I + M + 3-ways + FAMS	—
		D. Reported Behavior Change	
12. Condom (C)	RASEC	I + M + 3-ways + CERS	—
	GASEC	I + M + CGS	CS, CGS
13. Avoidhom (A)	RASEA	Saturated model	—
	GASEA	I + AS + AA + AG	AS

Table 2. (*Continued*)

Dependent Variable[a]	Model Tested[b] (1)	Best-Fitting Model ($p \geq .20$) (2)	Significant Interactions with Time (S)[c] (3)
14. Avoidmed (A)	RASEA	I + M + 3-ways + AAER + AARS	AS, AAS
	GASEA	I + AS + AG	AS

[a] See the appendix for exact question wording.

[b] Independent variables are identified by letter: gender = G, age = A, education = E, race = R, marital status = M, and study = S. The dependent variable is listed last in each model; its corresponding letter is provided in the left column (e.g., Homosex = H, Needles = N). Thus, the first model is the cross-classification of Homosex *by* Gender *by* Age *by* Study *by* Education.

[c] $p < .05$, two-tailed. Only the main effect of time and first-order interactions between demographics and time were tested.

measuring concern about AIDS becoming an epidemic for the public at large [Epidemic]; and the three behavior questions (avoiding homosexuals [Avoidhom], avoiding medical procedures involving a blood transfusion [Avoidmed], and using condoms [Condom]).[8] Of the 8 knowledge questions, only one—the belief that one could get AIDS by donating blood—showed an increase in inaccuracy over time;[9] responses to 5 questions showed gains in accuracy; and 2 (needle sharing and homosexual contact), no change.

All three of the questions about behavior—condom use, avoiding homosexuals, and avoiding medical procedures involving transfusions—showed statistically significant change in the sample as a whole, though condom use and avoiding homosexuals did so only in the models controlling for gender rather than race (see table 2, col. 3). It cannot be assumed that all of these changes are beneficial; for example, the increase in the percentage saying they would avoid homosexuals or places where homosexuals gather may be an indication of increasing prejudice against gays. It is also possible that a yea-saying response tendency was at work, so that respondents tended to give the same answer to all three questions about behavior, failing to distinguish among them, because they were asked one after another.

Nevertheless, the reported increase in condom use is the first indica-

---

8. In several of these (Sneezing, Work, Donate, Condom, and Avoidhom), however, only one of the two models tested showed the change as significant.

9. The significant change for Donate in the direction of greater inaccuracy emerges only when the effects of race, education, and age are modeled in addition to time.

tion in public opinion data of an increase in useful preventive behavior, and should not be minimized. It is supported by an analysis of condom sales between 1984 and 1988 by Moran et al. (1990), which indicates that there was a substantial increase during this period. Most of it, however (from 240.5 to 289.4 million), occurred between 1986 and 1987, before the government information campaign. Moran et al. note that the media rarely mentioned condoms before the Surgeon General's Report (1986), but did so frequently thereafter, with the number of newspaper items referring to condom use peaking in February, 1987, and then declining again. Thus, they conclude that the report, in conjunction with media attention, was effective in spurring condom sales. But such sales showed relatively little additional change (from 289.4 to 298.9 million) in response to the 1988 public information campaign.

Although there was no increase in personal concern about AIDS in the sample as a whole between 1987 and 1988, we thought that the increase in reported condom use might be attributable to increased concern among women, since only women reported an increase in condom use. However, there were no changes in concern among women between 1987 and 1988: 20.7% of women in 1987 and 19.2% in 1988 reported that they were very concerned about the possibility of contracting AIDS (data not shown).

Changes in Subgroup Knowledge, Attitudes, and Behavior

In order to assess changes in subgroups over time, we tested the significance of first-order interactions between each of the demographic variables and time (study). Relatively few such interactions were significant, indicating that little *differential* change in subgroup knowledge, opinions, or behavior had taken place during the year (table 2, col. 3).[10] The most frequent significant interactions were those between age and time and education and time. Interactions between race and time were significant for only three dependent variables (Donate, Sneeze, and Work), and marital status and gender were each involved in one significant interaction (Toilet and Condom, respectively).

When differential information gains occurred, they were not always in the direction of greater accuracy for groups less accurate to begin with. For example, nonwhites, who were somewhat less knowledgeable than whites in 1987, were significantly less knowledgeable in 1988 than they had been in 1987 on three of the eight transmission items (working next to others, sneezing, and donating blood) with other variables controlled (table 2, col. 3; see table 1, col. 3 for the trivariate relationships). And on two of the four transmission items on which

[10] For interactions, statistical significance is assessed using two-tailed tests.

education interacted significantly with time (needle sharing and donating blood), the better educated gained more than the less well educated (table 2, col. 3; see table 1, col. 4 for the trivariate relationships). Even in the case of age, two of the four items showing a significant interaction between age and time (donating blood and sneezing) indicated that the age group initially *more* knowledgeable had gained more in the course of the year (table 2, col. 3; see table 1, col. 2 for the trivariate relationships). Increases in inaccuracy were especially pronounced for the perception that one could get AIDS from donating blood, with nonwhites and those with less than a high school education both showing large changes for the worse.

As may be recalled, there was no discernible change in the sample as a whole so far as personal worry about AIDS was concerned. However, there was a significant interaction between personal worry and age (table 2, col. 3), with worry decreasing significantly among those under 45 (table 1, col. 2). On the other hand, as we have noted, concern about the possibility of AIDS becoming an epidemic for the population at large increased by 21 percentage points between 1987 and 1988 in the sample as a whole; but there was no differential change among subgroups.

As far as attitudes toward the regulation of others are concerned, only the interaction between age and time proved to be statistically significant (table 2, col. 3): Older people became more likely to agree that employers should have the right to fire workers who have AIDS, whereas younger people became somewhat less likely to agree with that statement (table 1, col. 2).

There was no clear pattern to subgroup changes in reported behavior. With respect to condom use, there was a significant first-order interaction between time and gender (table 2, col. 3), with women, who had been less likely to report condom use in 1987, showing more change than men between 1987 and 1988 (table 1, col. 6). On a second question about behavior—avoiding homosexuals—there were no differential subgroup changes over time (table 2, col. 3). And on the third question about behavior—avoiding medical procedures involving transfusions—there was again a significant interaction between time and age (table 2, col. 3): Older people showed the larger increase; younger people had reported higher levels of avoidance in 1987 and showed less change in the course of the year (table 1 col. 2).

Assessing the Causes of Change

As we have seen, some changes in knowledge, attitudes, and reported behavior did occur between 1987 and 1988. What evidence, if any, do we have that the Surgeon General's campaign, and in particular the brochure, "Understanding AIDS," was responsible for these changes?

Given the limitations of our research design, we cannot unambiguously attribute the changes we observe to the information campaign; other social changes may have been responsible instead, or in addition. It is also possible that the campaign had an effect even if we observe no change, since it is conceivable that, in the absence of the campaign, there would have been a negative change. Nevertheless, because we also have information about earlier and later changes in knowledge and attitudes, we can make reasonable inferences about the effects of the campaign itself.[11]

Earlier figures are available for 9 of the 14 questions replicated in the Gallup surveys of 1987 and 1988, although the wording of some of the questions varied slightly and most of the earlier surveys were done by organizations other than Gallup. Of the 8 knowledge questions, the 2 on sexual transmission were asked as a single question in prior surveys, so no direct comparison is possible. The other 6 showed no more change from 1987 to 1988 than they had prior to 1987 and, in most cases, less.

For four of these six questions, we were able to compare the Gallup figures with changes reported by ABC for a period of 18 months, ending 7 months before the 1987 Gallup survey. For the question about working next to someone, we used earlier data reported by CBS in September, 1985, and October, 1986, and for the question about donating blood, an NBC survey in January, 1986, and the Gallup survey in October, 1987. The questions, dates, and changes are shown in table 3.

Data made available to us by the Gallup Organization from a survey carried out in November, 1989,[12] further confirm that the 1988 figures were part of a trend toward increasing accuracy that had begun before 1987 and continued after 1988. Changes from 1988 to 1989 are shown in the third column of table 3. In 1989, the public gave increasingly accurate answers to questions about actual transmission routes and,

---

[11] It has been suggested that there is no reason to expect later changes to be as great as earlier ones, that past studies suggest a diffusion curve in which the rate of learning of new information declines after an initial period of rapid learning by those most attentive. But with the infusion of new information into the system and in the absence of ceiling effects, it seems reasonable to expect such changes. Furthermore, the fact that changes in the year after the campaign are as great as those during the year of the campaign strongly argues against the campaign itself being responsible.

[12] Like the two surveys analyzed here, this was a national, in-person survey carried out for the Gallup Poll. Only the questions about transmission and the two questions we had asked in 1988 about concern were included in 1989. For the November, 1989, survey, $N = 1{,}234$.

Table 3. Changes in Knowledge about AIDS Transmission

Knowledge about Transmission[b]	Percentage-Point Change[a] ABC (9/85–3/87)[c]	Gallup (10/87–10/88)	Gallup (10/88–11/89)
Needles	+1	−1	+1
Drink	−8	−6	−6
Sneezing	−16	−7	−6
Toilet	−12	−1	−11
Work	−11	−6	No change
Donate	−3	+1	+13

[a] For Needles, a plus means a gain in accuracy, and a minus means a loss in accuracy; for the other items, the reverse is true.
[b] See the appendix for exact question wording.
[c] For Work, change is measured by two CBS surveys, one in September, 1985, the other in October, 1986. For Donate, change is measured by a January, 1986, NBC survey and the Gallup October, 1987, survey.

with two exceptions, to questions about casual transmission as well. The dramatic exception was the question about donating blood, which showed a 13-point increase in inaccuracy between 1988 and 1989.[13] The other exception was the question about catching AIDS from working next to someone who has it, which showed no change.

Furthermore, the one transmission route not specifically mentioned in the Surgeon General's brochure—being sneezed on—showed change as great as any other item in the survey. And accuracy with respect to donating blood, which was specifically singled out for reassuring comment in the government brochure, actually deteriorated in the course of the campaign.[14] On the basis of all these findings, we conclude that the Surgeon General's campaign had minimal effects on the information level of the population at large.

This conclusion of minimal effects is supported by data reported by the National Center for Health Statistics (NCHS), which surveyed national samples of adults about AIDS every week from August to December, 1987, and again from May, 1988, until the present. Hardy and Dawson (1989) report that accuracy with respect to some questions about AIDS increased from May to December, 1988. But on questions having to do with misperceptions about how AIDS is transmitted, accuracy increased from May until July but then began to decline again, so that by December, 1988, "the proportion responding correctly to questions about casual transmission was often no different from that observed in May 1988" (1989, 3).

---

[13] Analyses indicate that this increase was in large part "real," although it was, in part, an artifact of question order. See Singer, Colasanto, and Rogers 1990.
[14] The British survey found a slight decline in the percentage believing one could get AIDS by donating blood, but the figures at the end of the information campaign were still 39% among adults and gays and 47% among youth. See Department of Health and Social Security and the Welsh Office 1987 (table 32).

We were able to make similar appraisals of the relative effect of the government information campaign on three other questions asked by Gallup in 1987 and 1988. Sentiment for firing workers declined between July and October of 1987, then stabilized between October, 1987, and October, 1988, the dates bracketing the information campaign. As we have seen, personal concern also remained stable between 1987 and 1988, and has, in fact, fluctuated only between 14% and 20% since Gallup first asked the question in August, 1985 (Singer, Rogers, and Corcoran 1987).

The question showing the largest change—more than 20 percentage points—during the period of the government information campaign was the one asking respondents whether they believed it likely that AIDS would become an epidemic for the population at large. (When Gallup repeated the question in 1989, the percentage responding "likely" remained at the 1988 level.) Although it is impossible to say whether the increase resulted from the government campaign alone, from media attention in combination with the campaign, or from still other causes, it is of interest that the British information campaign was likewise accompanied by an increase in general concern, from 48% to 65% (Department of Health and Social Security and the Welsh Office 1987, 69).[15] Therefore, the apparent increase in concern about the likelihood of an AIDS epidemic may be an unanticipated consequence of the information campaign.

As noted, the three questions asking about preventive behavior all showed small but statistically significant increases between 1987 and 1988. Unfortunately, these questions had been asked in a rather different form in 1986, and so no comparable earlier data are available against which to evaluate the 1987–88 change. Nor were they repeated by Gallup in 1989.

---

[15] It is easier to rule out the campaign as a source of change than to attribute observed change to the campaign. If changes in opinion during earlier and later time periods are as large as, or larger than, changes from before to after the campaign, we rule out the campaign as a source of change. But if there is a large change during the campaign not paralleled by earlier or later increases, we cannot be certain that the campaign was the sole agent responsible for the change. And as we have already noted, other polls put the 1987 level of concern at a higher level than Gallup.

Summary and Conclusions

During the period of a year, from October, 1987, to October, 1988, statistically significant changes occurred in the sample as a whole on 9 of 14 items on which comparisons between identically worded questions were possible. These 9 encompassed three of the four areas we investigated: knowledge, concern, and behavior. The only area showing no change was attitudes toward the regulation of others, measured in our data by a single question about employers' right to fire workers who have AIDS. The question we address in this section is the role of the government information campaign in bringing about these changes.

So far as changes in knowledge are concerned, we would argue that the campaign had no discernible effect. Those changes that occurred, though statistically significant, were smaller than those that had occurred during comparable periods in earlier years, even in the absence of a deliberate, nationwide public information campaign, and (with the exception of the belief that AIDS could be caught by working next to someone) were no larger than those that occurred between October, 1988, and November, 1989, well after the campaign had ended.

Nor was there any indication that subgroups who were less knowledgeable before the beginning of the campaign—for example, young people, singles, nonwhites, or the less well educated—gained more information as a result of the government's information campaign. Thus, there was no evidence that it succeeded in closing "knowledge gaps" between those who were better informed and those who were less well informed; indeed, on the basis of prior research (e.g., Cook et al. 1975; Gaziano 1983; Moore 1987; Tichenor, Donohue, and Olien 1970) we would not have expected it to have such effects.

The role of the campaign is perhaps most ambiguous with respect to beliefs about the riskiness of donating blood. Such beliefs, which had decreased before the beginning of the campaign, showed a slight but significant increase during the campaign and a very large increase in the year following. It is possible that the campaign delayed this increase, and if we had information about individual exposure to campaign materials, we would be better able to test this hypothesis.

In general, however, we believe that the changes in knowledge we observed resulted from a general climate of information, of which the government's campaign was one, but by no means the only or even the most important, element. Figures released by the Centers for Disease Control in 1990 indicate that news media coverage of AIDS/HIV infection peaked in 1987, with 11,852 news stories, up sharply from 5,377 in 1986 and 5,007 in 1985, and has declined again since then, to 7,574 news stories in 1988 and 5,231 in the first three quarters of 1989. Had the government campaign been launched earlier, at a time of sparser coverage, its impact might have been more noticeable. Coming when it did, its incremental effects appear to have been negligible, at least so far as changes in knowledge are concerned.

The government information campaign had no effect on attitudes toward the regulation of others, as far as we can tell on the basis of one item. This is perhaps not surprising, since there was no direct attempt to influence such attitudes during the campaign.

We know from other research (Price and Hsu 1989) that those who are more likely to believe that the AIDS virus can be transmitted by casual contact are also more likely to favor restrictive policies toward people who are infected, and so we might expect that a reduction in misperceptions would reduce the endorsement of restrictive policies as well. However, no such reduction is discernible in the single item included in our study. Clearly, this is an area that merits further research.

During the year bracketing the government information campaign, the reported use of condoms increased by a small but significant amount. Because the Surgeon General's brochure, as well as many of the spot ads developed in connection with the campaign, advocated the use of condoms as a way of reducing exposure to the AIDS virus, it is plausible to attribute the reported increase to the campaign.

The reported increase in condom use is paralleled by an increase in condom sales. But the analysis by Moran and his colleagues (1990) indicates that the largest increase in condom sales occurred between 1986 and 1987, before the beginning of the government's campaign but after the Surgeon General's Report on AIDS (1986), when media attention also peaked.

Future studies should continue to monitor trends in this important preventive behavior. Between 1985 and 1988, public opinion polls included fewer questions about behavior change than about any other aspect of AIDS. More items measuring behavior change are needed, as well as attention to reported reasons for it.

As we have noted, the largest change observed between 1987 and 1988 was a 20-percentage-point increase in concern that AIDS would become an epidemic for the population at large. Certainly, this message might reasonably have been inferred from the content of the Surgeon General's brochure, as well as from the fact that the campaign was targeted at the general population rather than those at special risk. In the absence of a comparison group, we hesitate to attribute this effect to the information campaign alone. But a comparable change in response to the British effort suggests that the campaign, perhaps in conjunction with news media coverage, may have been responsible.

The analyses reported in this article, together with our earlier analyses of trends in public opinion, clearly point to the importance of the mass media in bringing about change. Changes in knowledge before the campaign were as large as, and often larger than, changes during the campaign, at a time when the only source of information was the mass media. And though reported condom use appears to reflect the effects of the campaign, condom sales showed a greater increase in the preceding year, when news media coverage peaked. Undoubtedly, more could have been, and should be, done (e.g., Kinsella 1989; Shiltz 1987). But without the attention of the media, the public would be less aware, less well informed, and even less likely to engage in preventive behavior than appears to be the case.

Appendix

QUESTION WORDINGS FOR GALLUP POLLS, OCTOBER, 1987, AND OCTOBER, 1988

A. Knowledge about Transmission

As I read off each item, one at a time, would you tell me whether you think it is, or is not, a way for people to catch AIDS from someone who has it?

1.	Homosex	Intimate sexual contact with a person of the same sex? Yes; No/DK
2.	Hetersex	Intimate sexual contact with a person of the opposite sex? Yes; No/DK
3.	Needles	Sharing hypodermic needles? Yes; No/DK
4.	Toilet	From toilet seats? Yes/DK; No
5.	Drink	From a drinking glass? Yes/DK; No
6.	Donate	Donating blood? Yes/DK; No
7.	Sneezing	Being coughed or sneezed upon? Yes/DK; No
8.	Working	Working alongside or in close proximity to someone with AIDS? Yes/DK; No

B. Concern about Spread

9.	Epidemic	Do you think it is likely or is not likely that AIDS will eventually become an epidemic for the population at large? L, NL (DK excluded)
10.	Concerned	How concerned are you that you, yourself, will get AIDS—very concerned, a little concerned, not very concerned, or not at all concerned? VC; LC/NVC; NC (DK excluded)

C. Attitudes toward Regulation of Others

11.	Fire	Please tell me whether you agree or disagree with the following statement: Employers should have the right to dismiss an employee because that person has AIDS. Agree; Disagree/DK

D. Reported Behavior Change

Here is a list of some of the steps people are taking to reduce their chances of catching AIDS. Would you tell me which of these steps, if any, you yourself have taken. Just read off the number, or numbers:

12. Condom	Using condoms for the first time or using them more often than you used to. Yes; No/DK/NA
13. Avoidhom	Taking more care to avoid homosexual people or those you think are homosexual or places where homosexuals meet. Yes; No/DK/NA
14. Avoidmed	Avoiding any medical treatment that would involve receiving blood transfusions. Yes; No/DK/NA

References

Bishop, Yvonne, Stephen E. Fienberg, and Paul W. Holland. 1975. *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press.

Cook, Thomas D., Hilary Appleton, Ross F. Conner, Ann Shaffer, Gary Tamkin, and Stephen J. Weber. 1975. *"Sesame Street" Revisited*. New York: Russell Sage Foundation.

Department of Health and Social Security and the Welsh Office. 1987. *AIDS: Monitoring Response to the Public Education Campaign February 1986–February 1987*. London: Her Majesty's Stationery Office.

Gaziano, Cecilie. 1983. "The Knowledge Gap: An Analytical Review of Media Effects." *Communication Research* 10:447–86.

Goodman, Leo A. 1972. "A General Model for the Analysis of Surveys." *American Journal of Sociology* 77:1035–86.

Goodman, Leo A. 1978. *Analyzing Qualitative/Categorical Data: Log Linear Models and Latent Structure Analysis*, ed. Jay Magidson. Cambridge, MA: Abt.

Hardy, Ann M., and Deborah A. Dawson. 1989. "AIDS Knowledge and Attitudes for December 1988." *NCHS Advancedata* no. 175 (May 31).

Hyman, Herbert H., and Paul B. Sheatsley. 1947. "Some Reasons Why Information Campaigns Fail." *Public Opinion Quarterly* 11:412–23.

Kinsella, James. 1989. *In the Time of Plague*. New Brunswick, NJ: Rutgers University Press.

Mendelsohn, Harold. 1973. "Some Reasons Why Information Campaigns Can Succeed." *Public Opinion Quarterly* 37:50–61.

Moore, David W. 1987. "Political Campaigns and the Knowledge Gap Hypothesis." *Public Opinion Quarterly* 51:186–200.

Moran, John S., Harlan R. Janes, Thomas A. Peterman, and Katherine M. Stone. 1990. "Increase in Condom Sales following AIDS Education and Publicity, United States." *American Journal of Public Health* 80(5): 607–8.

Nelkin, Dorothy. 1987. "AIDS and the Social Sciences: Review of Useful Knowledge and Research Needs." *Review of Infectious Diseases* 9:980–86.

Price, Vincent, and Mei-Ling Hsu. 1989. "Factors Underlying Public Opinion on AIDS-related Issues: Predictors of Knowledge and Policy Preferences." Paper presented at the annual conference of the Midwest Association for Public Opinion Research, Chicago.

Rogers, Everett M., and J. Douglas Storey. 1987. "Communication Campaigns." In *Handbook of Communication Science*, ed. C. Berger and S. Chaffee. Newbury Park, CA: Sage.

Shiltz, Randy. 1987. *And the Band Played On*. New York: St. Martin's Press.

Singer, Eleanor, Diane Colasanto, and Theresa F. Rogers. 1990. "Context Effects in

Questions about AIDS." Paper presented at the annual meeting of the American Association for Public Opinion Research, Lancaster, PA.

Singer, Eleanor, Theresa F. Rogers, and Mary Corcoran. 1987. "The Polls—a Report: AIDS." *Public Opinion Quarterly* 51:580–95.

Singer, Eleanor, Theresa F. Rogers, and Grethe Lunde. 1988. "Trends in Knowledge, Attitudes, and Behavior in Response to AIDS among Subgroups of the General Population, 1985–87." Final report to the American Foundation for AIDS Research.

Surgeon General's Report on AIDS. 1986. Washington, DC: U.S. Department of Health and Human Services.

Tichenor, Philip J., George A. Donohue, and C. N. Olien. 1970. "Mass Media Flow and Differential Growth in Knowledge." *Public Opinion Quarterly* 34:159–70.

WOMEN CANDIDATES IN THE NEWS: AN EXAMINATION OF GENDER DIFFERENCES IN U.S. SENATE CAMPAIGN COVERAGE

KIM FRIDKIN KAHN and
EDIE N. GOLDENBERG

Abstract By covering male and female candidates differently, the news media may influence the success of female candidates for public office. A content analysis was conducted to assess potentially important differences in the newspaper coverage of a sample of male and female U.S. Senate candidates in the elections of 1982–86. The results of the study show that female candidates for the U.S. Senate are treated differently by the press. Female candidates receive less news coverage and the coverage they do receive concentrates more on their viability and less on their issue positions. Furthermore, female candidates' viability coverage is more negative than that of their male counterparts. Given these gender differences in press treatment, we would expect voters' recognition of male candidates to exceed that of female candidates and we would also expect evaluations of female candidates to be tied more closely to their perceived viability. Because female candidates are often considered noncompetitive by the press, this attention to the horserace may lead voters to develop more negative evaluations of female candidates. These results suggest that current patterns of press coverage may serve as a critical obstacle for women running for the U.S. Senate.

Women have always been underrepresented in the U.S. Senate. A total of sixteen women have served in the Senate and only six of these women have served full terms. The representation of women has not improved in recent years: two women serve in the U.S. Senate today,

KIM FRIDKIN KAHN is Assistant Professor at Arizona State University and EDIE N. GOLDENBERG is Dean of the College of Literature, Science, and the Arts and Professor of Political Science and Public Policy at the University of Michigan. The authors would like to thank the Horace H. Rackham School of Graduate Studies at the University of Michigan for support through the Research Partnership Program. The authors would also like to thank Amy Cuzzola, William Wehrle, and Susan Grantham for their help in data collection.

Public Opinion Quarterly 55:180–199 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-02/$2.50

the same number as in 1936. Scholars have examined a number of potential reasons for the lack of women in higher political office. Four principal explanations have emerged from this exploration and each of the four has received some empirical support: (1) women rarely seek political office (Deber 1982); (2) women often lack the necessary political resources (Epstein 1981; Gertzog 1979); (3) women often run in hopeless races (Bernstein 1986; Gertzog and Simard 1982); and (4) women are victims of sex stereotyping by voters (Boles and Durio 1980 and 1981; Bowman 1984; Sapiro 1982).

One explanation which has yet to receive any attention is that the news media may play a role in influencing the success of female candidates. Recent studies clarify the powerful role the news media can play in campaigns; the media can influence what voters learn about candidates as well as the criteria voters use when evaluating candidates (Brady and Johnston 1987; Erbring, Goldenberg, and Miller 1980; Iyengar and Kinder 1987). If the media rely on certain stereotypes when covering male and female candidates, and if this reliance creates differences in coverage, then media treatment can have important consequences for voter information and candidate preference. The news media, by covering male and female candidates differently, may hinder a woman's possibility of success in the political arena.

Our study examines systematic differences in the newspaper coverage of male and female U.S. Senate candidates. Senate races have been chosen because of the demonstrated importance of the news media in statewide races (Goldenberg and Traugott 1987). Coverage of male and female candidates may vary along a number of dimensions, and these differences in coverage may influence voters' evaluations of the candidates. We will investigate two broad types of coverage patterns: gender differences in the quantity of press coverage and gender differences in the substance of coverage. The possible electoral consequences of these various differences in coverage will also be addressed.

Study Design

The data reported here are drawn from a content analysis of newspaper coverage of 26 U.S. Senate races in 1984 and 1986.[1] These two elections were chosen in order to include as large a sample of Senate races with female candidates as possible, without extending back so many years that coverage patterns would no longer represent current practice.

---

[1] Only general election races between the two major party candidates have been included in the sample.

Before drawing the sample of Senate races for analysis, races were divided into the following seven types.

1. Male Incumbent vs. Female Challenger
2. Male Incumbent vs. Male Challenger
3. Female Incumbent vs. Male Challenger
4. Female Incumbent vs. Female Challenger
5. Male vs. Male in an Open Race
6. Male vs. Female in an Open Race
7. Female vs. Female in an Open Race

Each type of race was further divided by the level of competitiveness based on final vote returns. Competitive races were those races where the winner garnered no more than 55% of the vote, somewhat competitive races were races where the winner won more than 55% but less than 65% of the vote, and noncompetitive races were races where the winner received more than 65% of the vote. This categorization yielded 21 possible categories of races. For each of the categories, three races were randomly selected for analysis—if available.[2] The total number of races and the number of races chosen for the sample are shown in table 1. This sampling procedure yielded 26 Senate races—12 in 1984 and 12 in 1986. In addition, two races with female candidates from 1982 were included in the sample to bolster the number of competitive Senate contests with female candidates.

The newspaper with the largest circulation in each state in the sample was analyzed.[3] (See the appendix for a listing of the newspapers and the corresponding Senate races.) For the content analysis, every day from September 1 through the day after the election was coded; any item located anywhere in the newspaper that mentioned either candidate was coded.[4] This includes news articles, columns, editorials, letters to the editor, magazine articles, and political cartoons.

---

[2] For some types, three races for each level of competitiveness did not exist. In these instances, all available races were included.

[3] The newspapers in Texas and Colorado with the largest circulation were unavailable and others were substituted. In Texas, the largest newspaper was the *Houston Chronicle* with a circulation of 444,557; it was replaced by the *Houston Post* with a circulation of 326,556. In Colorado, the largest newspaper was the *Denver Rocky Mountain News* with a circulation of 317,638; it was replaced by the *Denver Post* with a circulation of 234,303.

[4] In those cases where the state primary was held after September 1, coding began the day after the primary. The following states held late primaries.

Colorado, Minnesota, New Hampshire, Wyoming	September 11, 1984
Massachusetts	September 18, 1984
Florida	September 2, 1986
Arizona, Maryland, New Hampshire	September 9, 1986

Table 1. A Comparison of the Population and Sample of Senate Races

Type of Race	Population	Sample
Male incumbent vs. female challenger		
Competitive	2[a]	2
Somewhat competitive[b]	1	1
Not competitive	9	3
Male incumbent vs. male challenger		
Competitive	15	3
Somewhat competitive	7	3
Not competitive	21	3
Female incumbent vs. male challenger		
Competitive	1	1
Somewhat competitive	0	0
Not competitive	1	1
Female incumbent vs. female challenger		
Competitive	0	0
Somewhat competitive	0	0
Not competitive	0	0
Male vs. male in an open race		
Competitive	5	3
Somewhat competitive	1	1
Not competitive	2	2
Male vs. female in an open race		
Competitive	2[a]	2
Somewhat competitive	0	0
Not competitive	0	0
Female vs. female in an open race		
Competitive	0	0
Somewhat competitive	0	0
Not competitive	1	1
Total	68	26

[a] Includes the competitive race with a female candidate from 1982.

[b] During the data analysis, the competitive and somewhat competitive races in the sample were combined.

The newspaper with the largest circulation was chosen for analysis because of its reach and potential impact on large numbers of people in the state. Yet, the largest papers are more professional and probably less likely than smaller papers to stereotype male and female candidates. If gender differences in coverage are found in these large papers, even more striking differences may exist in smaller papers.

NEWSPAPER COVERAGE OF MALE AND FEMALE SENATE CANDIDATES

Differences in newspaper coverage of male and female U.S. Senate candidates may be caused by reporters' and editors' reliance on standard criteria of newsworthiness (Epstein 1974; Graber 1989). For instance, because female candidates for the U.S. Senate are relatively rare, their candidacies may be treated as more newsworthy and they may subsequently receive more coverage than their male counterparts. Differences in the coverage of male and female candidates may also emerge if reporters and editors hold certain sex stereotypes and if these stereotypes influence coverage patterns. For instance, past research shows that voters believe that male candidates can deal more effectively with certain issues (i.e., the economy and foreign policy) while female candidates are better able to deal with such issues as education and health (Gallup 1984; National Women's Political Caucus 1987; Rosenwasser et al. 1987; Sapiro 1982). If reporters and editors hold these same stereotypes, then the substance of issue coverage for male and female candidates may vary accordingly.

If male and female candidates are covered differently by the press, this differential treatment can have important consequences for citizens' evaluations of the candidates and for eventual vote choice. First, variations in the quantity of male and female candidate coverage may lead to differences in voters' recognition of the candidates. If female candidates receive more attention in the press, then voters may have an easier time recognizing these candidates. This hypothesized advantage in recognition may be an important resource for female candidates since voters will rarely support a candidate they do not recognize (Goldenberg and Traugott 1984). Second, the substance of campaign coverage may vary for male and female candidates, and this difference in emphasis may influence voters' evaluations of the candidates. For example, the discussion of the horserace (i.e., who is ahead and who is behind in the public opinion polls) may differ for male and female candidates. Because female candidates have historically been less viable than male candidates, the press may spend more time discussing the horserace when covering female candidates. This preoccupation with the horserace may lead voters to become more concerned with the female candidates' electability, leading to more negative evaluations of female candidates (Iyengar and Kinder 1987). In examining gender differences in coverage of U.S. Senate candidates, we will examine both differences in the quantity of coverage and differences in the substance of coverage.

Quantity of Media Coverage

Because the amount of coverage a candidate receives is related to the voter's recognition of that candidate (Goldenberg and Traugott 1984 and 1987), gender differences in press attention can be consequential. As discussed above, standard definitions of news (Epstein 1974; Graber 1989) may lead reporters and editors to see female candidates as especially newsworthy and female candidates may therefore receive more attention in the news. On the other hand, if reporters and editors regard female candidates as less viable and thus less important than their male counterparts, then female candidates may receive less news exposure.

When we examine news coverage by type of race, we find that races with male candidates received more press attention than races with female candidates. As the data in table 2 indicate, 38.4 paragraphs a day were devoted to all-male Senate races while only 29.6 paragraphs a day were written about races that included female candidates.[5] If the all-male races are more competitive than the races that include female candidates, we would expect this difference in coverage. Both Clarke and Evans (1983) and Goldenberg and Traugott (1984), in their studies of House campaigns, find that competitive races receive substantially more press coverage than noncompetitive races. When we examine Senate coverage by level of competition,[6] we find that the male candidates' advantage in press attention persists—even when we control for the competitiveness of the race.[7] Another important and potentially

---

[5] Although measures of statistical significance are not entirely appropriate here because the sample is not a random sample from a larger population (i.e., the entire population has been examined for female incumbents and female candidates in open races), statistical significance is reported in the tables because they provide a rough guide for interpreting the importance of the reported differences. In addition, because our sample of races is necessarily small, we will treat as probably meaningful some differences that do not reach conventional levels of significance.

[6] For data analysis, races were initially divided into three levels of competitiveness based on (1) preelection ratings published in the *Congressional Quarterly* Special Election Issue and (2) final vote returns. Because coverage patterns for competitive and somewhat competitive races are similar, these two categories of races are combined for analysis.

[7] Although the differences here are not statistically significant, they are substantively interesting. It should be noted that the small sample size here makes it difficult to obtain statistically significant differences.

Table 2. Average Number of Paragraphs Published Each Day

	Total[a]	*N*[b]	*p*-Value[c]
All races			
All-male races	38.4 (13.5)	15	$p < .10$
Races that include female candidate(s)	29.6 (12.3)	11	
Competitive races			
All-male races	41.3 (13.3)	10	n.s.
Races that include female candidate(s)	36.5 (12.9)	6	
Noncompetitive races			
All-male races	30.3 (11.9)	5	n.s.
Races that include female candidate(s)	21.3 (7.1)	5	

[a] Entries are the mean number of paragraphs published each day with the standard deviation in the parentheses.

[b] Entries are the number of races included for each type of race.

[c] The *p*-value is based on the *t*-statistic; n.s. = not significant.

confounding variable is the size of the state's newspaper. Newspapers vary in size, and if larger newspapers are more prevalent in the all-male races, then the greater number of stories in the all-male races could be explained by differences in newspaper size rather than by the gender of the candidates. However, newspaper size turns out not to be a significant factor: the average newshole for the all-male races is 4,094 column-inches while the average newshole for races that include female candidates is 3,939 column-inches, nearly as large.

Candidate coverage, like race coverage, may vary with the gender of the candidate. Given the findings discussed above, we would expect female candidates to capture less news attention than male candidates. This is exactly what we find: male candidates consistently receive more coverage than their female counterparts. Overall, an average of 12.9 paragraphs are published about male candidates each day while only 10.5 paragraphs are devoted to female candidates each day. Although the gender differences in coverage are not large or statistically significant, they are consistent and occur for both competitive and noncompetitive candidates.

These results suggest that voters will have a more difficult time acquiring information about female candidates. This lack of available information, because it will likely inhibit voter recognition of female candidates, may limit a woman's chances of election, since voters will rarely vote for a candidate they do not recognize.

Substance of Media Coverage

Besides examining gender differences in the quantity of coverage, it may be useful to explore gender differences in the substance of coverage. If female candidates are covered differently in the news, then these differences in reporting may influence the content of voters' evaluations of the candidates. Differences in the substance of coverage may be due to organizational incentives operating in news rooms or they may be due to pervasive sex stereotypes.

HORSERACE COVERAGE

Recent media research and work in the field of voting behavior suggest the importance of three types of campaign coverage: horserace, issue, and trait coverage. First, media studies of presidential campaign coverage show that the press spend most of their time discussing horserace aspects of campaigning. The discussion of the horserace, at the expense of issue coverage, occurs in both television and newspaper coverage of presidential elections (Graber 1989; Patterson 1980; Robinson and Sheehan 1983).

This emphasis on the horserace aspects of political campaigns seems to have electoral consequences. First, Iyengar and Kinder's (1987) experimental work on agenda setting and priming indicate that the news media, by concentrating on the horserace, may make the candidates' viability more salient to voters. Second, research on presidential primary voting by Bartels (1987) and Brady and Johnston (1987) suggests that the substance of the horserace coverage, by influencing voters' assessments of the candidates' viability, can influence voters' overall evaluations of the candidates. Therefore, if voters are taking cues from the media with regard to the viability of the various candidates, then this media information may influence voters' attitudes toward the candidates.

Quantity of Horserace Coverage. Given the potential importance of horserace coverage, it is important to examine whether there are systematic differences in the amount and substance of horserace coverage for male and female candidates. Because female Senate candidates have traditionally been less viable than their male counterparts, the press may spend more time discussing the viability of female candidates. As the data in table 3 indicate, horserace coverage is more

Table 3. Percentage of Articles Devoted to the Horserace

	All[a]	*N*[b]	*p*-Value[c]
Gender of candidate			
Male candidates	21	5,708	$p < .01$
Female candidates	27	1,378	
Gender of candidate by competitiveness of race			
Male in competitive race	22	4,271	$p < .05$
Female in competitive race	27	847	
Male in noncompetitive race	17	1,437	$p < .01$
Female in noncompetitive race	27	531	
Gender of candidate by candidate status			
Male incumbent	21	1,954	$p < .01$
Female incumbent	30	254	
Male challenger	19	1,606	$p < .01$
Female challenger	28	602	
Male in an open race	21	2,148	n.s.
Female in an open race	24	522	

NOTE: Horserace coverage includes any discussion of a candidate's chances of winning, including comparisons of the two candidates' organizations, discussion of poll results, and discussion of the candidates' performance in campaign debates.

[a] Entries are the percentage of articles devoted to the horserace.

[b] Entries are the number of articles coded for each candidate type.

[c] The *p*-value is based on the difference in proportions test. This particular test examines differences between the proportion of articles devoted to the horserace for male and female candidates.

prevalent for female candidates. Overall, 21% of the articles written about male candidates discuss the horserace, while 27% of the articles written about female candidates do. This greater emphasis on the horserace for female candidates occurs for all types of candidates: competitive and noncompetitive candidates, incumbents, challengers, and candidates in open races.

Viability Assessments by the Press. Because the emphasis on the horserace aspects of campaigns is consistently greater for female Senate candidates, voters who are exposed to this information may become more concerned with the viability of female candidates. This could lead to more negative evaluations of female candidates *if* the horserace information about female candidates is negative. To examine the content of the horserace coverage, every press assessment of a candidate's viability was rated on a scale from 1 to 4, with 1 indicating "sure loser," 2 indicating "somewhat competitive," 3 indicating "competitive," and 4 indicating "likely winner." As the data in table 4 show, the press give female candidates more negative viability as-

Table 4. Press Assessments of Candidate Viability

	All[a]	N[b]	p-Value[c]
Gender of candidate			
Male candidates	3.1 (.64)	971	$p < .01$
Female candidates	2.5 (.88)	289	
Gender of candidate by candidate status			
Male incumbent	3.5 (.42)	351	$p < .01$
Female incumbent	2.7 (.67)	49	
Male challenger	2.4 (.61)	252	$p < .01$
Female challenger	2.1 (.61)	128	
Male in an open race	2.9 (.45)	368	n.s.
Female in an open race	3.0 (.46)	112	

NOTE: Press viability assessments were rated on a four-point scale where 4 denotes "likely winner," 3 denotes "competitive," 2 denotes "somewhat competitive," and 1 denotes "sure loser."

[a] Entries are mean viability assessments by the press with the standard deviation in parentheses.

[b] Entries are the number of articles that mention the candidates' viability.

[c] The p-value is based on the t-statistic.

sessments. Male candidates, on average, are seen as competitive (3.1) while female candidates are labeled somewhat less competitive (2.5).

This difference in viability assessments for male and female candidates may be a reflection of status differences for male and female candidates. In both the population of Senate races and the sample of Senate races used here, female candidates are more likely than male candidates to be challengers. In our sample, 50% of the female candidates are challengers while only 27% of the male candidates are challengers.[8] These gender differences in status suggest that female candidates may be treated by the press as less viable because they are more likely to be challengers. Yet, when we control for the status of the candidate, we find that female incumbents and female challengers continue to receive more negative viability assessments than their male counterparts (see table 4). Female candidates in open races, on the other hand, are not disadvantaged by press assessments of viability. Since only a handful of female candidates ran in open races, these results suggest that, in most cases, the press consider female candidates to be less viable than their male counterparts.[9]

---

[8] This difference in the proportion of male and female challengers echoes the gender differences in status found in the population of Senate races: 73% of all female candidates running for the Senate between 1984 and 1986 were challengers while only 38% of the male candidates were challengers.

Campaign Resources. News attention to campaign resources (e.g., the strength of the candidate's campaign organization, the candidate's fundraising capabilities) was also examined, since such discussion is likely to contribute to voters' judgments of a candidate's viability. The conventional wisdom suggests that female candidates are often disadvantaged with regard to campaign financing (Burrel 1985; Lake 1984; Mandel 1981). If reporters hold these same preconceptions, or if female candidates do indeed have fewer resources, then reporters may concentrate on the female candidates' lack of campaign resources in their press coverage. In this study, we find that negative resources are discussed more frequently for female candidates. Only 5% of the articles written about male candidates mention a lack of resources, while more than 10% of the articles about female candidates do. These differences remain even when we examine news coverage by status: negative resources are emphasized more in the coverage of female incumbents, female challengers, and females in open races than in the coverage of their male counterparts.[10]

These results show that female candidates often receive more horserace coverage than comparable male candidates, and this coverage is often negative in terms of the discussion of the candidates' viability and the candidates' campaign organization. This does not bode well for female Senate candidates seeking election. First, the news media's emphasis on horserace issues for female candidates may make the female candidates' viability more salient for voters. Second, because the content of the horserace coverage is often more negative for female candidates, voters may develop negative evaluations of female candidates and these evaluations may influence vote choice.

ISSUE COVERAGE

Media scholars have generally bemoaned the lack of attention given to issues in the coverage of candidates for electoral office. This lack of issue coverage is widely documented at the presidential level (e.g., Graber 1972; Patterson 1980; Robinson and Sheehan 1983; Russonello and Wolf 1979). For instance, Russonello and Wolf (1979), in their study of newspaper coverage of the 1968 and 1976 presidential campaigns, find that only 24% of the campaign coverage focused on issues while the horserace was discussed in almost half (49%) of the campaign articles. Because the public depends on the news media for information about candidates, the lack of issue information in the press may make it difficult for people to make policy evaluations of the candidates running for office.

---

[9] Only 3 of the 16 women running for the U.S. Senate between 1984 and 1986 were candidates in open races.

[10] These gender differences in the negative coverage of resources are statistically significant at $p < .05$.

Issue voting does occur in Senate elections when voters are given adequate information regarding the stands of the opposing candidates. For instance, both Abramowitz (1981) and Wright and Berkman (1986) find that, when there is a clear distinction between opposing candidates, voters take issue and ideological information into consideration when they make their vote choice. Since issue information can be consequential, it is important to see whether issue coverage in campaigns is different for male and female candidates.

When we examine the amount of campaign coverage devoted to issues each day, we find that more than 3.5 paragraphs are published about male candidates' issue stands each day while less than 3.0 issue paragraphs are published each day for female candidates. The data in table 5 show that the gender differences in issue coverage occur for all types of candidates—incumbents, challengers, and candidates in open races.[11] Female candidates consistently receive less issue attention in the press than their male counterparts. This lack of issue attention in the press, coupled with the emphasis on the horserace, may lead voters to discount issues and emphasize viability when evaluating female candidates. Issues may play a more central role in the evaluations of male candidates, since issues receive more play and the horserace less play in the news coverage of their candidacies.

The substance of issue coverage, as well as the amount of coverage, may differ for male and female candidates. Research on sex stereotyping of political candidates clearly shows that voters believe that male candidates can deal more effectively with certain issues, while female candidates are more competent in other policy areas (e.g., Rosenwasser et al. 1987; Sapiro 1982). If reporters and editors hold these same stereotypes, then the issue coverage for male and female candidates may vary accordingly. Similarly, candidates themselves may stress different types of issues in their campaigns, and these differences in issue emphasis may influence coverage patterns.

[11] Although female incumbents, challengers, and open race candidates consistently receive less issue coverage than their male counterparts, these differences fail to reach statistical significance ($p < .10$).

Table 5. Average Number of Paragraphs Published about Issues Each Day

	All[a]	N[b]	p-Value[c]
Gender of candidate			
Male candidates	3.7 (2.1)	40	$p < .10$
Female candidates	2.9 (1.7)	12	
Type of candidate			
Male incumbent	3.7 (1.8)	15	n.s.
Female incumbent	3.3 (3.4)	2	
Male challenger	3.3 (2.0)	11	n.s.
Female challenger	2.3 (1.1)	6	
Male in an open race	4.1 (2.4)	14	n.s.
Female in an open race	3.6 (1.9)	4	

[a] Entries are the mean number of paragraphs published about issues each day with the standard deviation in parentheses.

[b] Entries are the number of candidates included for each type of race.

[c] The p-value is based on the t-statistic.

Based on past stereotyping research, issues were divided into "male" and "female" issues (Gallup 1984; National Women's Political Caucus 1987; Rosenwasser et al. 1987; Sapiro 1982). "Male" issues include those issues where men are seen as being more competent (e.g., foreign policy, defense, economics, and agriculture); "female" issues are those where women are considered more capable (e.g., minority rights, environment, abortion, school prayer, drugs, and discussions of social programs). As the data in table 6 indicate, "male" issue coverage is much more prevalent than "female" issue coverage, but the type of coverage varies with the gender of the candidate. Overall, "female" issues are discussed more frequently for female candidates (40%) than for male candidates (30%). Similarly, the presence of a woman in the race seems to encourage greater discussion of "female" issues—even for male candidates.

Without examining campaign materials, we cannot know whether different issues in the news reflect different emphases by candidates

Table 6. Percentage of "Male" and "Female" Issue Coverage

	Male Issues	Female Issues	N^a	p-Value
Gender of candidate				
Male candidates	70	30	9,327	$p < .01^b$
Female candidates	60	40	2,264	
Type of race				
Male candidates in all-male races	71	29	7,060	$p < .01^c$
Male candidates in mixed-gender races	66	34	2,267	
Female candidates	60	40	2,264	

NOTE: "Male" issues include foreign policy, defense spending, arms control, foreign trade, farm issues, and the economy; "female" issues include day-care, helping the poor, education, health care, women's rights, drug abuse, and the environment.

[a] Entries are the number of male and female issue paragraphs coded for each candidate type.

[b] The p-value is based on the difference in proportions test. This particular test examines differences in the proportion of paragraphs devoted to "male" issues and "female" issues for male and female candidates.

[c] The p-value is based on the difference in proportions test. The test here examines differences in the proportion of paragraphs devoted to "male" issues and "female" issues for the three types of candidates. Issue coverage for each of the three candidate types is significantly different from the others ($p < .01$).

or different news selection decisions by editors and reporters. We suspect a bit of both. We can begin to examine news selection decisions by seeing whether female reporters are more likely to cover races with women candidates and whether these reporters are more likely to report on "female" issues. We find that female reporters are more likely than male reporters to cover races with female candidates. In particular, female reporters cover mixed-gender races 47% of the time, while male reporters cover mixed-gender races only 30% of the time. Furthermore, female reporters discuss "female" issues more often than male reporters in their campaign coverage (35% vs. 28%), and this difference in issue emphasis is most pronounced in the coverage of female candidates. In fact, when covering female candidates, female reporters devote 42% of their issue discussion to "female" issues, while male reporters just spend 19% of their time covering "female" issues. These results suggest that the gender of the reporter may offer part of the explanation for the difference in issue emphasis for male and female candidates.[12]

[12] We also investigated whether the gender of the reporter influenced coverage of the horserace for male and female candidates. We found no significant reporter differences here.

These differences in the substance of issue coverage for men and women candidates can lead to differences in voters' evaluations of the candidates. Since media coverage can influence the salience of issues for the public (Erbring, Goldenberg, and Miller 1980; Iyengar and Kinder 1987; MacKuen 1981), voters exposed to races with female candidates may come to believe that "female" issues are important and these "female" issues, because they are more salient, may be used when voters evaluate the candidates (Iyengar and Kinder 1987; Iyengar, Peters, and Kinder 1982). The greater salience of these "female" issues, coupled with voters' sex stereotypes, may lead to more positive evaluations of the female candidates. For example, if education (a "female" issue) is emphasized in the campaign coverage of a race with a female candidate, voters exposed to the campaign coverage may come to believe that education is an important issue, and voters may evaluate the opposing candidates with regard to how well they can deal with educational problems. Since most voters believe that female candidates are better able to deal with educational issues (Gallup 1984; National Women's Political Caucus 1987; Rosenwasser et al. 1987; Sapiro 1982), the salience of education as a campaign issue may lead voters to develop more positive evaluations of the female Senate candidate.

TRAIT COVERAGE

Along with horserace and issue coverage, the discussion of personality traits is an important aspect of campaign coverage. The coverage of personality traits is prevalent in campaigns (Graber 1989; Patterson 1980; Robinson and Sheehan 1983), and recent voting studies suggest that voters' assessments of candidate traits can influence overall candidate evaluations (Abelson et al. 1982; Goldenberg, Traugott, and Kahn 1988; Markus, 1982). We wondered whether the coverage of candidate traits is more prevalent for female Senate candidates because female candidates are still somewhat novel and their personal characteristics may be especially interesting to reporters (and readers). If this is the case, trait-based evaluations may be more common for female candidates. Trait coverage, however, is no different for male and female candidates. In terms of the percentage of articles that mention candidate traits, we find that traits are discussed with equal frequency for female candidates (14%) and male candidates (13%).

Although we find no gender differences in the amount of trait coverage, the substance of trait coverage may differ for male and female candidates. Based on research on sex stereotypes, we can distinguish between "male" and "female" traits (Ashmore and del Boca 1979;

Boles and Durio 1980 and 1981; Ruble and Ruble 1982). "Male" traits are those traits that are consistently associated with men (e.g., independent, objective, competitive, strong leader, insensitive, aggressive, unemotional, ambitious, and tough) while "female" traits are linked with women (e.g., dependent, noncompetitive, passive, gentle, emotional, weak leader, and compassionate). Just as with issues, we might expect reporters' coverage of candidates' personality traits to reflect these sex stereotypes. If the substance of trait coverage in the news is different for male and female candidates, then these differences may lead to different evaluations of the candidates. In particular, if different traits are emphasized in the news, certain traits may become more salient or central to voters' evaluations of male candidates (e.g., independence) while other traits (e.g., compassion) may be more central to voters' evaluations of female candidates. As the data in table 7 show, the discussion of "male" traits is more prevalent in campaign coverage, but discussion of "female" traits is more common in the coverage of female candidates than male candidates.

As with issues, the sex of the author continues to make a difference, with female authors appearing to be more sensitive to the gender of the candidate. Female authors devote over 80% of their trait discussion to "male" traits when covering male candidates, while "male" traits are discussed only 52% of the time for female candidates. Male authors seem to be "gender-blind" in their trait coverage, devoting about 60% of their trait coverage to "male" traits, regardless of the gender of the candidate. As with the substance of issue coverage, the female reporters' predilection to discuss "female" traits more frequently for women candidates may lead to more positive evaluations of the women candidates. The coverage of "female" traits by female reporters may lead voters to believe these traits are more important, and voters may think about these personality traits when evaluating the candidates. Given voters' sex stereotypes and the salience of the "female" traits, voters may evaluate female candidates more positively than male candidates. These results suggest that the presence of female reporters, and the accompanying coverage pattern, may encourage more positive evaluations of female candidates. Yet, female reporters are relatively rare; only 24% of all the campaign articles in the present sample were written by women.

Conclusions

Female candidates for the U.S. Senate do not receive the same press treatment as male candidates. The results of the content analysis show

Table 7. Percentage of "Male" and "Female" Trait Coverage

	Male Traits	Female Traits	N[a]	p-Value[b]
Gender of candidate				
Male candidates	75	25	462	$p < .01$
Female candidates	60	40	174	
Sex of author				
Male author				
Male candidates	64	36	383	n.s.
Female candidates	63	37	72	
Female author				
Male candidates	81	19	59	$p < .01$
Female candidates	52	48	46	

NOTE: "Male" traits include the following: independent, objective, competitive, strong leader, insensitive, aggressive, unemotional, ambitious, and tough. "Female" traits include the following: dependent, noncompetitive, passive, gentle, emotional, weak leader, and compassionate.

[a] Entries are the number of male and female trait paragraphs coded for each candidate type.

[b] The p-value is based on the difference in proportions test. This particular test examines differences in the proportion of paragraphs devoted to "male" traits and "female" traits for male and female candidates.

that female candidates receive less news coverage and the coverage they do receive concentrates more on their viability and less on their issue positions. Furthermore, female candidates' viability coverage is more negative than that of their male counterparts. Given these gender differences in press treatment, we would expect voters' recognition of male candidates to exceed that of female candidates, and we would also expect evaluations of female candidates to be tied more closely to their perceived viability. Because female candidates are often considered noncompetitive by the press, this attention to the horserace may lead voters to develop more negative evaluations of female candidates. These results suggest that current patterns of press coverage may serve as a critical obstacle for women running for the U.S. Senate.

On the positive side, the substance of both issue and personality trait coverage is influenced by the gender of the reporter, with female reporters placing greater emphasis on "female" issues and traits, especially in their coverage of female candidates. This emphasis on female issues and traits, coupled with voters' sex stereotypes, may encourage more positive evaluations of the female candidates. If this is the case, the presence of female reporters on the staff of major metropolitan newspapers may be an important resource for female candidates running for electoral office.

Appendix: Sample of Races

Male Incumbent vs. Female Challenger

1. Colorado (1984)	*Denver Post*
2. Illinois (1986)	*Chicago Tribune*
3. Indiana (1986)	*Indianapolis Star*
4. Minnesota (1984)	*Minneapolis Star and Tribune*
5. Missouri (1982)	*Kansas City Times*
6. Nebraska (1984)	*Omaha World-Herald*

Female Incumbent vs. Male Challenger

1. Florida (1986)	*Miami Herald*
2. Kansas (1984)	*Wichita Eagle-Beacon*

Male vs. Female in an Open Race

1. Missouri (1986)	*St. Louis Post-Dispatch*
2. New Jersey (1982)	*Newark Star-Ledger*

Female vs. Female in an Open Race

1. Maryland (1986)	*Baltimore Sun*

Male Incumbent vs. Male Challenger

1. California (1986)	*Los Angeles Times*
2. Illinois (1984)	*Chicago Tribune*
3. Iowa (1984)	*Des Moines Register*
4. Iowa (1986)	*Des Moines Register*
5. Michigan (1984)	*Detroit News*
6. New Hampshire (1984)	*Manchester Union Leader*
7. New Hampshire (1986)	*Manchester Union Leader*
8. Pennsylvania (1986)	*Philadelphia Inquirer*
9. Wyoming (1984)	*Casper Star-Tribune*

Male vs. Male in an Open Race

1. Arizona (1986)	*Arizona Republic*
2. Colorado (1986)	*Denver Post*
3. Massachusetts (1984)	*Boston Globe*
4. Nevada (1986)	*Las Vegas Review Journal*
5. Tennessee (1984)	*Memphis Commercial Appeal*
6. Texas (1984)	*Houston Post*

References

Abelson, R. P., D. R. Kinder, M. D. Peters, and S. T. Fiske. 1982. "Affective and Semantic Components in Political Person Perception." *Journal of Personality and Social Psychology* 42:619–30.

Abramowitz, A. 1981. "Choices and Echoes in the 1978 U.S. Senate Elections: A Research Note." *American Journal of Political Science* 25:112–18.

Ashmore, R., and F. del Boca. 1979. "Sex Stereotypes and Implicit Personality Theory: Toward a Cognitive-Social Psychological Conceptualization." *Sex Roles* 5:219–48.

Bartels, L. M. 1987. "Candidate Choice and the Dynamics of the Presidential Nominating Process." *American Journal of Political Science* 31:1–30.

Bernstein, R. 1986. "Why Are There So Few Women in the House?" *Western Political Quarterly* 39:155–64.

Boles, J., and H. Durio. 1980. "Social Stereotyping of Males and Females in Elected Office: The Implications of an Attitudinal Study." Paper presented at the annual meeting of the Midwest Political Science Association, Chicago, IL.

Boles J., and H. Durio. 1981. "Political Woman and Superwoman: Sex Stereotyping of Females in Elected Office." Paper presented at the annual meeting of the Midwest Political Science Association, Chicago, IL.

Bowman, A. 1984. "Physical Attractiveness and Electability: Looks and Votes." *Women and Politics* 4:55–65.

Brady, H. E., and R. Johnston. 1987. "What's the Primary Message: Horserace or Issue Journalism." In *Media and Momentum*, ed. G. R. Orren and N. W. Polsby, 127–86. Chatham, NJ: Chatham House.

Burrel, B. 1985. "Women's and Men's Campaigns for the U.S. House of Representatives, 1972–1982: A Finance Gap?" *American Politics Quarterly* 13:251–72.

Clarke, P., and S. H. Evans. 1983. *Covering Campaigns: Journalism in Congressional Elections*. Stanford: Stanford University Press.

Deber, R. B. 1982. "'The Fault, Dear Brutus': Women as Congressional Candidates in Pennsylvania." *Journal of Politics* 144:463–79.

Epstein, C. F. 1981. "Women and Power: The Role of Women in Politics in the United States." In *Access to Power: Cross-National Studies of Women and Elites*, ed. C. F. Epstein and R. L. Closer, 124–46. Boston: George Allen and Unwin.

Epstein, L. J. 1974. *News from Nowhere*. New York: Vintage Books.

Erbring, L., E. N. Goldenberg, and A. H. Miller. 1980. "Front-Page News and Real World Cues: A New Look at Agenda Setting by the Media." *American Journal of Political Science* 24:16–49.

Gallup Report. 1984. Vol. 228:2–14.

Gertzog, I. 1979. "Changing Patterns of Female Recruitment to the U.S. House of Representatives." *Legislative Studies Quarterly* 4:429–45.

Gertzog, I., and M. M. Simard. 1982. "Women and 'Hopeless' Congressional Candidacies: Nomination Frequency 1916–1978." *American Politics Quarterly* 9:449–66.

Goldenberg, E. N., and M. W. Traugott. 1984. *Campaigning for Congress*. Washington, DC: Congressional Quarterly Press.

Goldenberg, E. N., and M. W. Traugott. 1987. "Mass Media Effects on Recognizing and Rating Candidates in U.S. Senate Elections." In *Campaigns in the News: Mass Media and Congressional Elections*, ed. J. P. Vermeer, 109–31. New York: Greenwood Press.

Goldenberg, E. N., M. W. Traugott, and K. F. Kahn. 1988. "Voter Assessments of Presidential and Senatorial Candidates." Paper presented at the annual meeting of the Midwest Political Science Association, Chicago, IL.

Graber, D. A. 1972. "Personal Qualities in Presidential Images: The Contribution of the Press." *Midwest Journal of Political Science* 16:46–76.

Graber, D. A. 1989. *Mass Media and American Politics*. Washington, DC: Congressional Quarterly Press.

Iyengar, S., and D. R. Kinder. 1987. *News That Matters*. Chicago: University of Chicago Press.

Iyengar, S., M. D. Peters, and D. R. Kinder. 1982. "Experimental Demonstrations of the 'Not So Minimal' Consequences of Television News Programs." *American Political Science Review* 76:848–58.

Lake, C. 1984. "Impact of Gender on Campaigns: A Study of Men and Women Candidates in 1982." Prepared for the National Women's Political Caucus, Washington, DC.

MacKuen, M. 1981. "Social Communication and the Mass Policy Agenda." In *More than News*, ed. Michael B. MacKuen and Steven Lane Coombs, 19–46. Beverly Hills: Sage.

Mandel, R. 1981. *In the Running*. New Haven: Ticknor and Fields.

Markus, G. B. 1982. "Political Attitudes during an Election Year: A Report of the 1980 NES Panel Study." *American Political Science Review* 75:538–60.

National Women's Political Caucus Survey (NWPC). 1987. Washington, DC: NWPC.

Patterson, T. E. 1980. *The Mass Media Election*. New York: Praeger.

Robinson, M. J., and M. A. Sheehan. 1983. *Over the Wire and on TV: CBS and UPI in Campaign '80*. New York: Sage.

Rosenwasser, S. M., R. Rogers, S. Fling, K. Silvers-Pickens, and J. Butemeyer. 1987. "Attitudes toward Women and Men in Politics: Perceived Male and Female Candidate Competencies and Participant Personality Characteristics." *Political Psychology* 8:191–200.

Ruble, D. N., and T. L. Ruble. 1982. "Sex Stereotypes." In *In the Eye of the Beholder: Contemporary Issues in Stereotyping*, ed. A. Miller, 188–252. New York: Praeger.

Russonello, J., and F. Wolf. 1979. "Newspaper Coverage of the 1976 and 1968 Presidential Campaigns." *Journalism Quarterly* 56:360–65.

Sapiro, V. 1982. "If U.S. Senator Baker Were a Woman: An Experimental Study of Candidate Images." *Political Psychology* 2:61–83.

Wright, G. C., Jr., and M. B. Berkman. 1986. "Candidates and Policy in U.S. Senate Elections." *American Political Science Review* 80:569–88.

THE ANSWERING MACHINE POSES MANY QUESTIONS FOR TELEPHONE SURVEY RESEARCHERS

PETER S. TUCKEL and
BARRY M. FEINBERG

Abstract The proliferation of the telephone answering machine raises a number of questions about the continued efficiency of the telephone as a data-gathering mechanism. The most critical of these is what effect the use of these machines might have on establishing contact with potential respondents. If individuals routinely use these machines to screen calls, then their accessibility to telephone survey researchers will be restricted. A second question is the degree to which respondents who own answering machines and are reachable are likely to participate in a survey. A third question is the extent to which the incidence of the answering machine as a response disposition might vary by the time of calling and the respondent's place of residence. The findings of this study, based on a nationwide survey, point to a significant proportion of answering machine owners being reachable and willing to participate. Furthermore, the answering machine appears to be in use more on weekends than on weekday evenings and in more urbanized areas than in areas with fewer inhabitants.

Just as the use of the telephone ushered in a new era in the administration of large-scale surveys, recent advances in telephone technology raise a number of questions about the continued efficiency of the tele-

PETER S. TUCKEL is Associate Professor of Sociology at Hunter College, C.U.N.Y. BARRY M. FEINBERG is Vice President, Research, at Burson-Marsteller. At the time the research for this article was conducted, Barry Feinberg was Vice President and Senior Research Executive at Audits & Surveys. The authors wish to thank Richard Maisel, Department of Sociology, New York University; Manfred Kuechler, Department of Sociology, Hunter College, C.U.N.Y.; Dennis Hodgson, Department of Sociology, Fairfield University; and Elaine Leppo of In Vision, Inc., for their helpful comments on this paper. The authors are also grateful to Bob Ruppe, Elizabeth Lou, and Alan Perry of Audits & Surveys for the technical assistance they lent to this project. An earlier version of this paper was presented at the 1989 annual meeting of the American Association for Public Opinion Research.

Public Opinion Quarterly 55:200–217 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-02/$2.50

phone as a data-gathering mechanism. These advances in technology include: the telephone answering machine, the mobile phone, and a device to identify callers.

This study examines the effect of the answering machine on accessing respondents in telephone surveys. Surprisingly, little scholarly attention has been accorded this topic. The already widespread and increasing penetration of the answering machine into American households argues for the need to subject this phenomenon to systematic analysis. The research discussed in this paper is designed to be a first step toward meeting this need.

Proliferation of the Telephone Answering Machine

Estimates of the penetration level of telephone answering machines into American households (as of the spring of 1988) vary within a fairly narrow range, from 20% to 26%. The figure at the lower end of this range is provided by the Electronic Industries Association (EIA 1989) and is based on manufacturers' sales to wholesale distributors and certain retail outlets. This figure is probably biased downward for two reasons. First, some of the manufacturers might view volume of sales as being proprietary and, therefore, might be reluctant to disclose this information to the EIA. Second, there is a lag between the time the data are initially gathered and the time they are disseminated.

Somewhat higher and almost identical estimates of the ownership of telephone answering machines are found in two other studies based on survey data. One study, conducted by Walker Research (1988), reports that 25% of telephone households own an answering machine. Importantly, this study also finds that half of the owners acknowledge use of the machines to "screen their calls at least some of the time."

A second survey-based study, undertaken by the Yankee Group (1988), finds that 26% of telephone households possess an answering machine. The Yankee Group study further identifies subgroups in the population whose ownership of answering machines exceeds the average. Among those groups that are overrepresented as owners are: income earners over $50,000 (49.3%), home-based entrepreneurs (45.7%), individuals serviced by the Pacific Telesis Regional Bell Holding Company (41.5%), college graduates (40.0%), singles (33.3%), individuals 25–44 years old (32.6%), and men (31.5%).

Since the Walker Research and the Yankee Group studies themselves were administered over the telephone, it is likely they, too, underestimate the actual incidence of answering machine ownership, although not to the same degree as the EIA estimate.[1] As reported, some respondents screen their calls. This practice would serve to depress estimates of answering machine ownership based upon verbal responses in a telephone survey.

Not only do a significant proportion of American households currently possess an answering machine, but indications are that ownership will continue to rise in the near future. The Yankee Group study reports that the penetration of answering machines into American households increased by approximately 15% within the span of less than three years. With the marked reduction in the purchase price of answering machines, the study predicts that the ownership of these machines will continue to spiral upward.

The proliferation of the telephone answering machine poses a number of important questions bearing on the conduct of telephone survey research. The most important of these questions is: How accessible are answering machine owners to telephone survey researchers? If individuals routinely use these machines to screen calls, then, obviously, contact cannot be initiated by the survey researcher.

A second question is the degree to which respondents who own answering machines and are reachable are likely to participate in a survey. Are these individuals more inclined than others to assent to being interviewed?

A third question concerns the variability in answering machine disposition as a function of the time of calling and the geographic location of the respondent's household. Is answering machine disposition more common, for example, if calls are made on weekday evenings versus weekends, or if respondents live in densely populated areas versus areas with fewer inhabitants?

This study addresses each of these questions. The next section describes the data set and our study's methodology.

Data and Methodology

The data base for this study was derived from one wave (May, 1988) of a continuous tracking study in the automotive area carried out by Audits & Surveys.

For the host study, a national random digit dial sample of 4,884 telephone numbers was generated by Audits & Survey's sampling department and placed on the CATI queuing system that automatically assigns numbers to interviewers and dials the numbers. The selection of the sample of telephone numbers followed a general set of procedures for constructing a national probability sample with one modification (Audits & Surveys 1987). The general set of procedures consists of drawing a proportionate stratified sample of exchanges from all U.S. exchanges first stratified by census region and then by metropolitan and nonmetropolitan areas within each region. Next, following the Waksberg method (1978), "working" banks of 1,000 consecutive numbers are randomly chosen within each of the selected exchanges.

---

[1] Readers interested in obtaining more detailed information about the surveys conducted by Walker Research and the Yankee Group can contact them at the following addresses: Walker Research, 3939 Priority Way South Drive, Indianapolis, IN 46280-0432; The Yankee Group, 200 Portland Street, Boston, MA 02114.

In this particular study, however, a subset of "working" banks that yielded a high incidence of residential telephone numbers based on previous experience with all of the banks was identified. The sample of telephone numbers was then chosen from this subset of "working" banks.

This final modification represents a departure from a probability sample. One bias that might intrude as a result of this modification of the sample design is that telephone households in large metropolitan areas might be underrepresented in the sample. The reason for this potential bias is that geographic areas in which there are concentrations of commercial enterprises (namely, large metropolitan areas) might be associated with a lower frequency of "working" banks that have a high incidence of residential telephone numbers. A comparison of the regional distribution of telephone numbers in the sample with the regional distribution of telephone households in the U.S. population suggests, however, that if this bias exists it is of a small magnitude.[2]

Interviewers were instructed to make an initial and up to two callbacks to complete an interview with a potential respondent. In a few cases, more than three dials were made to complete an interview.

The calling period extended from May 2 to June 3, 1988. In general, the final disposition for a given household number was arrived at within a one-week period. The calling schedule was set up so that calls were rotated, using a cycle of two separate weekdays and one weekend day. In the case of numbers that produced a busy response, the calling

---

2. To measure the possible bias that might result from using "working" banks with a high incidence of residential telephone numbers, we undertook a three-stage analysis. First, we assigned each of the telephone numbers in our sample to one of four groups categorized according to the population of the county in which a given telephone number is located. In dividing the counties into four groups by population, we relied upon the A. C. Nielson Company's (1989) classification scheme, which rank orders the counties from the most populated (size A) to the least inhabited (size D).

Next, we computed the frequency of telephone numbers in the entire sample that fell into each of the four county size-groups. Finally, we compared this distribution to that of an estimate of all U.S. telephone households according to county size-group (Simmons Market Research Bureau 1989). In none of the four categories did the difference exceed 2.2 percentage points. It seems safe to conclude, therefore, that telephone numbers in large metropolitan areas are not underrepresented in our sample.

schedule departed from this pattern. For these numbers, additional call-attempts could be initiated as long as a minimum of 20 minutes separated each of the dialings.

With business and nonworking numbers (identified at the conclusion of the calling) excluded, the sample consisted of 2,981 telephone numbers.

All individuals 18 years old or older qualified to be interviewed. Each qualified individual who agreed to participate in the survey was administered a screening questionnaire. If an individual met additional criteria, he or she also was administered a full-length questionnaire. Included among the items in the full-length questionnaire was a battery of demographic questions. Altogether, there were 1,061 qualified individuals who completed the interview, of whom 300 completed the longer questionnaire.

The identification of households with answering machines was accomplished in two ways. First, whenever an interviewer encountered an answering machine, this was recorded on CATI using a specially created disposition code. (Henceforth, this shall be referred to as the answering machine response disposition.) No message was left on the machine concerning the nature of the call. Second, at the conclusion of *either* the screening questionnaire or the full-length questionnaire, a very brief sequence of questions was asked in which the respondent indicated whether the household had an answering machine normally connected to the phone number on which the interview was being conducted.[3]

After the fieldwork was completed, the CATI response disposition data file was combined with the data file for the completed interviews in which respondents reported ownership or nonownership of an answering machine.

This study design has some limitations that should be noted here. First, as mentioned earlier, in most instances households were called up to a maximum of three times before final disposition. To measure more precisely the effect that the answering machine has on establishing contact with potential respondents, it would have been preferable to increase the maximum number of call-attempts. Nevertheless, it should be pointed out that it is a common practice among many telephone survey researchers, particularly those operating in the commercial domain, to limit the number of call-attempts to three. Furthermore, there is widespread use, by academicians and others, of the results of commercially generated surveys that have been carried out under conditions similar to those employed here.

---

[3] The exact wording of the question about answering machine ownership was: "Do you have a telephone answering machine normally connected to the telephone number we are talking on now?"

A second limitation attached to this study is that demographic data were gathered on only a subset ($N = 300$) of the total number of respondents who completed the interview. This small number precluded constructing a profile of answering machine owners and making comparisons between the characteristics of owners versus nonowners.

Accessibility of Respondents with Answering Machines

Two competing hypotheses might be posited concerning the accessibility of respondents with answering machines. The first hypothesis might be that these respondents would be less susceptible to being contacted by survey researchers than other respondents. Underlying this hypothesis is the notion that answering machine owners might have purchased their machines with the express purpose of avoiding unwanted calls. It is possible that many of these individuals share a similar orientation toward the outside world—a world that they perceive as being increasingly intrusive and from which they seek psychological shelter. In marketing parlance, this group of individuals might be thought of as the *cocooners*—those who attempt to insulate themselves as much as possible from the public domain.[4] We shall refer to this first hypothesis as the cocooning hypothesis.

The alternative hypothesis might be that answering machine owners would be more accessible to telephone survey researchers than respondents without these machines. The reasoning underlying this hypothesis is that the answering machine owner, more than the typical respondent, might feel a heightened need to be connected with the outside world. Rather than obstructing communications with others, the answering machine functions in just the opposite way for this type of respondent—providing a means of maximizing accessibility to outsiders. Accordingly, this type of respondent would tend to utilize his or her answering machine only when he or she is not at home. We might label this group of respondents the *connectors* and, as in our first hypothesis, assign this name to our second hypothesis.

To measure the degree of accessibility of answering machine owners, we undertook two separate analyses. In the first, we examined the difference between the observed frequency of an answering machine as a response disposition and the frequency we would expect if all an-

---

4. *Cocooners* comes from the term *cocooning*, which originated with Faith Popcorn, of Brain Reserve, a New York City–based marketing company.

swering machine owners were to use their machines to the maximum extent. In the second analysis, we investigated the overall accessibility of answering machine owners and compared their accessibility with that of nonowners.

THE INCIDENCE OF THE ANSWERING MACHINE AS A RESPONSE DISPOSITION

According to the two survey-based studies cited previously, approximately 25% of all telephone households possess an answering machine. This figure could be thought of as roughly approximating the maximum percentage of residential telephone numbers to which an answering machine might be attached.[5] It is clear that, on a given call-attempt, the incidence of having an answering machine as a response disposition would be less than this maximum percentage. First, we can safely assume that not all homes with answering machines use them on a continual basis. Second, the utilization of answering machines might be reduced precisely during those hours in which telephone surveyors are most likely to conduct interviews. The question, then, becomes how wide a gap exists between the actual and potential frequency of occurrence of having an answering machine as a response disposition.

Table 1 displays the final response disposition of each of the sampled telephone numbers. The final response disposition refers to the status of each household number at the conclusion of the calling. To clarify the meaning of this term, let us take the answering machine final response disposition as an example. In this response category, the term refers to a number that, at the end of the calling, has: (1) not been categorized as business, nonworking, ineligible, partial, refusal, or completed interview, and (2) on at least one previous call-attempt produced an answering machine recording. In the case of a number that produced two response categories (e.g., answering machine and call-back), however, the final response disposition depends upon the temporal sequence of these two response categories. If a number yielded an answering machine on one call-attempt followed by a call-

---

[5] The 25% figure is not exact because it incorporates two biases. Although the biases operate in different directions and, therefore, might cancel each other out, their net effect cannot be precisely determined. On the one hand, as we pointed out earlier, the incidence of telephone households with answering machines probably exceeds 25%. This is because this figure was derived from studies that were themselves based on telephone interviews and in which a certain proportion of respondents affirmed their use of answering machines to screen their calls. On the other hand, the number of residential *telephone lines* is slightly greater than the number of *telephone households* (see Groves and Kahn 1979). Despite the lack of precision attached to the 25% figure, we believe it can still serve as a rough estimate of the maximum potential number of times an answering machine might be linked to a telephone line.

Table 1. Final Response Disposition of Sampled Telephone Numbers

Disposition	Total Frequency (%)	Household Frequency[a] (%)
Nonworking	28.3	—
Business	10.7	—
No answer	12.0	19.8
Busy	1.3	2.1
Answering machine[b]	3.4	5.7
Household refusal	11.9	19.6
Respondent refusal	1.8	3.0
Ineligible[c]	0.4	0.7
Call-back	8.0	13.3
Partial[d]	0.2	0.3
Completed[e]	21.6	35.6
Total[f]	99.6	100.1
N	4,884	2,981

[a] Known business and nonworking numbers are excluded.

[b] Answering machine = household numbers that produced an answering machine recording.

[c] Ineligible = households that did not include a respondent 18 years old or older.

[d] Partial = interviews that were terminated by the respondent before completion.

[e] Completed = household numbers that yielded a completed interview, including interviewees who said they owned an answering machine.

[f] Columns do not total 100% due to rounding.

back on a subsequent call-attempt, the final response disposition would be labeled a "call-back"; if this pattern were reversed, the final response disposition would be "answering machine."

Table 1 presents the proportion of the final response disposition for both total numbers (including business and nonworking numbers) and household numbers only. Inspection of table 1 shows that only 5.7% of the household telephone numbers resulted in an answering machine disposition at the conclusion of the calling. Though this figure is modest, it could be masking a higher incidence of answering machine dispositions, especially if there were numerous instances in which an answering machine contact was followed by a call-back.

The frequency distribution of the response categories at the time of each of the survey's first three call-attempts are shown in table 2.[6] As the number of call-attempts progresses from 1 to 3, the numeric base shrinks because the refusals, ineligibles, partials, and completed interviews from the preceding call(s) are removed. Known business and nonworking numbers are omitted from the data base.[7] For ease of presentation, certain response categories are grouped together. Table 2 reveals that the answering machine category constitutes 6.7% of all the calls made at the time of the first call-attempt, 6.4% made at the time of the second call-attempt, and 9.2% made at the time of the third call-attempt. While these percentages are not insubstantial, it is clear that a wide disparity exists between the incidence of answering machines in homes (not forgetting the inexactitude surrounding this estimate) and the occurrence of having an answering machine as a response category on a given call-attempt.

Table 2 also shows the percentage of calls at the time of each call-attempt that comprise completed interviews in which the respondent states he or she owns an answering machine, states he or she does not own an answering machine, or refuses to disclose information about ownership status. The figures for completed interviews in which ownership is affirmed range from 3.9% at the time of call-attempt 1 to 1.5% at the time of the third call-attempt. Another way of looking at these figures is that 19.3% of all the completed interviews at the end of the first three call-attempts of the survey in which answering machine status was established were conducted with respondents who reported owning an answering machine.

These findings suggest that a significant proportion of answering machine owners are accessible to telephone surveyors. If the estimates of the number of telephones equipped with answering machines in the population are accurate, then a marked discrepancy exists between

---

[6] The data base for this and subsequent analyses in this study is restricted to responses obtained on the first three call-attempts. The reasons for imposing this restriction are: First, only a small fraction of the telephone numbers in the sample were dialed more than three times (3.9%) and, of these, close to two-fifths turned out to be nonworking numbers. Second, the numbers that were dialed more than three times were not randomly distributed among the previous noncontacts. Of the numbers called additional times (including those that attained nonworking status), 38% had a busy response on the third call-attempt. In any event, the additional calls netted only 12 working numbers associated with an answering machine disposition.

[7] Two investigations carried out by Audits & Surveys on the response rates in other studies revealed that 15% and 30%, respectively, of the "no answers" are, in fact, business numbers. If we consider these percentages to be separate estimates of the concealed rate of business numbers within the "no answer" category and apply these estimates to the data provided in table 2, the percentage frequency of the answering machine disposition (based on residential numbers) increases somewhat at the time of each call-attempt. The adjusted frequencies would increase by a magnitude of approximately 0.4% to 0.9% at the time of each of the first two call-attempts and by a magnitude of 0.7% to 1.6% at the time of the third call-attempt.

Table 2. Response Disposition by Call-Attempt (in percentages)

Disposition	Call 1	Call 2	Call 3
No answer or busy	45.5	53.2	58.7
Answering machine	6.7	6.4	9.2
Call-back, ineligible, or partial	14.1	15.9	15.8
Refusal	13.1	9.6	6.7
Completed	20.6	14.9	9.7
With self-report of answering machine	(3.9)	(2.6)	(1.5)
With self-report of no answering machine	(15.5)	(11.1)	(7.4)
Answering machine status unknown	(1.2)	(1.2)	(0.8)
N	2,981	1,982	1,486

NOTE: Known business and nonworking numbers (identified at the conclusion of the calling) are excluded.

the incidence of answering machine households and the incidence of an answering machine as a response disposition on a given call-attempt. Furthermore, fully one-fifth of the completed interviews in the present study (in which answering machine status was ascertained) were conducted with respondents who affirmed they owned an answering machine.

THE ACCESSIBILITY AND COOPERATIVENESS OF ANSWERING MACHINE OWNERS

In addition to the preceding analysis, we also investigated the overall likelihood of contacting an answering machine household and the relative likelihood of contacting an answering machine household versus other households. To measure the overall likelihood, we isolated household numbers that produced an answering machine recording on either the first, second, or third call-attempt. We then divided this subpopulation into three categories: (1) households that were contacted, but in which a completed interview was not secured (i.e., refusals, ineligibles, call-backs, and partials); (2) households in which a completed interview was secured; and (3) households that were never contacted. The results displayed in table 3 reveal that a sizable proportion of answering machine households (55.9%) can be reached during two call-attempts. Moreover, one-fifth of these households produce completed interviews during these two call-attempts.

Table 3. Accessibility of Households with Answering Machines

	Answering Machine Households
Contacted, no interview	36.1%
Contacted, completed interview	19.8
Never contacted	44.1
N	288

NOTE: Known business and nonworking numbers (identified at the conclusion of the calling) are excluded.

Of related interest is how the accessibility of answering machine households compares to that of other households. In carrying out this comparative analysis, we first divided the noncontacts at the end of the first call-attempt into three groups: those with a response disposition of an answering machine (group 1), those with a response disposition of no answer (group 2), and those with a response disposition of busy (group 3). Second, we calculated the proportion of each of these three groups whom we were able to contact over the course of the next two call-attempts. These contacted households included the ineligibles, call-backs, refusals, partials, and completors.

Numbers that generated an answering machine disposition on call-attempts 2 or 3 were retained in the analysis. If this were not done, an asymmetry would result in the computation of the contact rates for the different groups. For example, an answering machine disposition for the no answer group (group 2) would not be counted as a noncontact, whereas a no answer disposition for the answering machine group (group 1) would be counted as a noncontact.

By confining the analysis to the noncontacts at the end of the first call-attempt and observing their response patterns over the next two call-attempts we standardize, in effect, the number of opportunities respondents in each of the groups have to be contacted.[8]

---

[8] An alternative design for this analysis would be to classify all household numbers by their answering machine status. According to this design, households that produced an answering machine recording or a completed interview with self-reported ownership of an answering machine on *any* of the first three call-attempts would be designated as "answering machine households"; all others would be designated as "non–answering machine households." Next, the contact rate for these two types of households would be computed.

Adoption of this strategy, however, would bias the contact rate of the non–answering

Table 4. Contact Rate at the End of the Third Call-Attempt by Type of Response at the Time of the First Call-Attempt

	Response Group		
	Answering Machine	No Answer	Busy
Contact	56.8%	46.8%	54.3%
Noncontact	43.2	53.2	45.7
N	199	1,172	184

NOTE: Known business and nonworking numbers (identified at the conclusion of the calling) are excluded.

The contact rates of the three groups are presented in table 4. The difference in the values of the rates between two of the three groups is noticeable. A substantially larger proportion of the answering machine group were contacted (56.8%) than were contacted in the no answer group (47.0%; $z = +2.63$, $p < .01$, two-tailed). Only a slight difference, though, exists between the contact rate of the answering machine group and the busy group—2.5 percentage points.

The reader will recall that two investigations of nonresponse in other studies revealed that 15% and 30%, respectively, of the no answers were concealed business numbers. To eliminate the bias that might be introduced into the analysis by their presence, we calculated contact rates for the no answer group adjusted at a lower and higher level. The no answer group that is adjusted at the lower level assumes that 15% of the no answer responses that recur at both call-attempts 2 and 3 are really business numbers and, therefore, are removed from the ranks of the noncontacts for this group. Similarly, the no answer group adjusted at the higher level assumes that 30% of the no answer responses that recur at both call-attempts 2 and 3 are really business numbers and, thus, this proportion is removed from the noncontacts for this group.

The contact rates adjusted for the two levels of the no answer group are presented in table 5. Importantly, the rates of both adjusted no

---

machine households upward. For example, households that yielded response dispositions of refusal, ineligible, or partial on the first call-attempt (comprising 13.6% of all the working numbers) would be assigned to the non–answering machine household group and be classified as contacts. Yet the answering machine status of these households is indeterminate, as we have no way of confirming that they do not possess answering machines.

Table 5. Contact Rate at the End of the Third Call-Attempt by Type of No Answer at the Time of the First Call-Attempt

	No Answer Group	
	Adjusted at Low Level	Adjusted at High Level
Contact	50.1%	54.0%
Noncontact	49.9	46.0
N	1,093	1,014

answer groups are still lower than the corresponding rate for the answering machine group. However, only the pairwise comparison between the no answer group adjusted at the lower level and the answering machine group is statistically significant and, then, only if a relatively high significance level is employed ($z = +1.75$, $p < .10$, two-tailed).

Beyond measuring the varying degree of accessibility of the answering machine, no answer, and busy groups, we also examined the extent to which respondents in these groups, once contacted, would be willing to participate in the survey. To accomplish this we calculated the proportion of the refusals and the completors combined in each of the three groups who completed the interview by the end of the third call-attempt.

Table 6 displays the results of this cross-tabulation and reveals modest differences in the completion rate between the answering machine group and the rate for the no answer and busy groups. Approximately 67% of the answering machine group agree to participate in the survey, compared to a corresponding figure of about 60% for the two other groups. The difference in the completion rates between the answering machine group and the no answer and busy groups are not statistically significant. Nevertheless, the direction of the differences in the rates suggest that answering machine owners might be more disposed toward being interviewed than members of the other two groups.

Taken together, the data contained in tables 4 through 6 could be construed as lending greater support to the connectors hypothesis than the cocooning hypothesis. It seems that answering machine owners (classified on the basis of an answering machine recording at the time of the first call-attempt) are more amenable to being contacted and

Table 6. Completion Rate at the End of the Third Call-Attempt by Type of Response at the Time of the First Call-Attempt

	Response Group		
	Answering Machine	No Answer	Busy
Completed interview	67.1%	60.5%	60.0%
Refused interview	32.9	39.5	40.0
N	73	382	75

interviewed than other individuals who initially fall into the category of noncontacts.

The Effect of Time and Place on the Incidence of the Answering Machine as a Response Disposition

This section focuses on the variability in the incidence of an answering machine response disposition associated with the time of calling and the respondent's place of residence. The analysis is divided into two phases—one concerned with the effects of the time of calling, the other with the effects of the geographic location of the respondent's household.

To carry out the first phase of the analysis, we grouped the times calls were made into two categories: weekday evenings (Monday through Friday, 5:01 P.M. to midnight) and weekends (Saturday and Sunday). We then cross-tabulated the frequency of response dispositions by these two time intervals, controlling for the number of the call-attempt. This control was imposed on the data because, if the call-attempts were aggregated, the distribution of these call-attempts would be skewed toward the hard-to-contact cases which, in turn, might be differentially associated with the timing of the calls.

The data in table 7 show that, for each call-attempt, the answering machine disposition is a more common occurrence on weekends than on weekday evenings. For the second and the third call-attempts, the differences are highly significant ($z = +4.32$, $p < .001$, two-tailed for call-attempt 2; $z = +2.66$, $p < .01$, two-tailed for call-attempt 3). However, for the first call-attempt, the difference does not prove to be statistically significant.

Table 7. Frequency of the Answering Machine Response Disposition on Weekday Evenings vs. Weekends

Response Disposition	Call 1		Call 2		Call 3	
	Weekday Evenings	Weekend	Weekday Evenings	Weekend	Weekday Evenings	Weekend
Answering machine	6.3%	8.0%	5.2%	16.3%	8.0%	13.6%
All others	93.7	92.0	94.8	83.7	92.0	86.4
N	2,372	600	1,764	215	1,174	308

NOTE: Known business and nonworking numbers (identified at the conclusion of the calling) are excluded. Calls made during weekday day hours ($N = 19$) are also excluded.

The greater frequency of the answering machine disposition on weekends might be attributed to the tendency of respondents to be away from home more on the weekends than on weekday evenings.

In the second phase of this analysis, we examined the varying likelihood of obtaining an answering machine disposition by the population size of the county in which the respondent's household was located. To measure population size, we again used the A. C. Nielson Company's (1989) classification scheme (discussed in note 2).

Table 8 displays the frequency of the response dispositions by the four county size–groups controlling for the number of call-attempts. The reason for controlling for the number of call-attempts is essentially the same as that used in the first phase of this analysis: if the response dispositions were summed over the three call-attempts, the resulting distribution would be weighted toward the hard-to-contact cases which might be differentially related to the county size–group.

An inspection of table 8 reveals, for each call-attempt, a positive relationship between the frequency of the answering machine disposition and the population size of the county in which respondents live. In each case, the relationship is statistically significant (for call attempts 1, 2, and 3, respectively: $\gamma = +.25$, $p < .01$; $\gamma = +.37$, $p < .001$; $\gamma = +.40$, $p < .001$).

That the answering machine disposition is more prevalent in more populous areas is not surprising, given the demographic profile of answering machine owners discussed above. Those individuals who are single, more educated, or upper income tend to be clustered in counties with denser populations.

Table 8. Frequency of the Answering Machine Response Disposition for Different County Size–Groups (in percentages)

Response Disposition	Call 1				Call 2				Call 3			
	A	B	C	D	A	B	C	D	A	B	C	D
Answering machine	8.7	5.7	6.0	2.8	9.4	4.2	4.6	2.5	13.3	7.0	5.7	2.6
All others	91.3	94.3	94.0	97.2	90.6	95.8	95.4	97.5	86.7	93.0	94.3	97.4
N	1,323	929	367	324	895	621	241	201	679	458	176	155

NOTE: Known business and nonworking numbers (identified at the conclusion of the calling) are excluded. Telephone numbers not assigned a county size–group ($N = 60$) are also excluded. Counties grouped by population according to A. C. Nielson 1989.

Conclusions

The major finding to emerge from this study is that a substantial proportion of answering machine households are accessible to telephone survey researchers. This finding is based on four specific observations. First, the incidence of an answering machine disposition on a given call-attempt falls well below the estimated national average of households with an answering machine. Second, approximately one-fifth of the *completed* interviews in our study were carried out with respondents who reported they owned an answering machine. Third, among households that yielded an answering machine disposition on any of the call-attempts, a sizable proportion (56%) resulted in a contact during the two remaining call-attempts. Fourth, among households that initially produced either an answering machine, no answer, or busy response, the subsequent contact rate for the answering machine households was greater than the rate for the no answer households and roughly the same as the busy households.

This study also provides evidence, although not as clear-cut, that individuals in households with an answering machine disposition on the first call-attempt display a greater propensity to participate in a survey than those in households with a no answer or busy response on the first call-attempt.

The study further shows that the answering machine disposition is more prevalent on weekends than on weekday evenings. One recommendation that might flow from this finding is to schedule call-backs with households found to have an answering machine on a weekday evening rather than on a weekend.

It is also clear from the study that the use of answering machines is greater in more urban areas of the country. This finding is consistent with the pattern of diffusion of electronic technology in our society, in which urban areas serve as points of origination.

Clearly, the present study represents only a first step toward understanding the role played by the answering machine in conducting telephone survey research. Periodic studies should be undertaken to monitor any change in the effect of the answering machine on telephone surveying, both as the number of answering machine households in society expands and as the nature of telephone technology evolves. The findings of our study are based on those who currently own an answering machine—an estimated 25% of telephone households. As the answering machine continues to proliferate, it is likely that the sociodemographic characteristics of owners will change. Accompanying this change may well be different orientations toward communicating with the outside world and, correspondingly, different behavior patterns associated with the utilization of the answering machine.

Additionally, the nature of telephone technology is changing rapidly. In the foreseeable future, for example, answering machines will have the capability to bar entry to originating telephone numbers that are not "recognized." This new capability will permit individuals to use their telephones differently. Thus, technological innovations, such as the caller-identification feature, may also induce profound change in the way the answering machine is currently used. Both the anticipated change in the demographic composition of answering machine owners and developments in telephone technology support the need for continued research in this area.

References

A. C. Nielson. 1989. *DMA Market Profiles: 1989 Nielson Station Index*. Chicago, IL: A. C. Nielson.

Audits & Surveys. 1987. "Audits and Surveys National Telephone Sample." Mimeo.

EIA. 1989. "Telephone Conversation." Electronic Industries Association, May 30.

Groves, R. M., and R. L. Kahn. 1979. *Surveys by Telephone: A National Comparison with Personal Interviews*. New York: Academic Press.

Simmons Market Research Bureau. 1989. "1989 Study of Media and Markets." Special computer run. New York: Simmons Market Research.

Waksberg, J. 1978. "Sampling Methods for Random-Digit Dialing." *Journal of the American Statistical Association* 73:40–46.

Walker Research. 1988. *Industry Image Study*. 8th ed. Indianapolis, IN: Walker Research.

Yankee Group. 1988. "Telephone Answering Devices: Gaining Consumer Acceptance." *Yankeevision* (August): 13–20.

THE ELECTORATE'S PARTISAN EVALUATIONS: EVIDENCE OF A CONTINUING DEMOCRATIC EDGE

JOHN G. GEER

Abstract Scholars have devoted a good deal of attention to studying changes in how the public evaluates the Democratic and Republican parties. However, there have been few attempts to examine the underlying components of the changes in these overall evaluations of the parties in detail. By recoding the Center for Political Studies' open-ended likes/dislikes questions for parties, this paper maps change in the underlying partisan evaluations of the electorate since 1952. The results suggest that the Democrats have remained the favored party, despite some gains by the GOP in recent years. Of the eight issues studied, the Republicans have made significant inroads only on the economic front. The findings also highlight the obstacles that face the Republican's effort to gain majority status, suggesting why it has remained the minority party for over 50 years, and why it is likely to remain in that position in the near future.

With the defeat of Michael Dukakis in 1988, the Democrats have lost five of the last six presidential elections, winning ten or fewer states in four of those contests. The Democrats have also lost a sizable share of self-identified partisans in recent years. In 1980, for instance, the Democrats held a twenty percentage point lead over the Republicans. By 1988, that lead had shrunk to about six percentage points. In addition, there are indications of a pro-Republican tide among the youth of the nation, suggesting that further gains by the GOP may be on the horizon. Work by Cavanagh and Sundquist (1985), Norpoth (1987), Norpoth and Kagay (1989), Petrocik (1987), Shanks and Miller (1989), and Wattenberg (1987 and 1990) all lend support to the claim that the fortunes of the Democratic party have been waning.

JOHN G. GEER teaches at Arizona State University. The author thanks Thad Brown, Pat Crittenden, Kim Kahn, Pat Kenney, Bob Luskin, Warren Miller, Jim Stimson, and Tom Rochon for their helpful comments on previous drafts of this paper. A much earlier version of this paper was presented at the 1988 WPSA meetings in San Francisco. Data used in this paper were made available by the ICPSR. The author bears sole responsibility for the analyses and interpretations presented here.

Public Opinion Quarterly 55:218–231 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-02/$2.50

While the GOP has made gains in the 1980s, this trend need not continue. The fortunes of the parties have had peaks and valleys over the last 50 years. In 1964, for instance, the GOP appeared to be retreating on nearly all fronts. Yet by 1968, the GOP had won the White House and had also made sizable inroads into the Democrat's edge in self-identified partisans. Such shifts suggest that we should be cautious when interpreting the recent changes in the public's attitudes toward the parties. Caution may be particularly appropriate in this case, given that the Democrats retain a plurality of self-identified partisans, control both houses of Congress, and hold a majority of governorships and legislatures in the fifty states.

To detect the depth of the GOP's gains and thereby develop a better sense of the relative standing of the two parties, I examine changes in the public's evaluations of Democratic and Republican parties since 1952. This examination does *not* include the public's assessment of presidential contenders. By setting aside attitudes toward the two nominees, one avoids contaminating the estimates of partisan support in the electorate. It is possible, for instance, that the gains by the GOP in recent years are tied, in part, to Reagan's personal popularity rather than any enduring change in how the public evaluates the two parties. Thus, by limiting attention to just the public's evaluation of the parties, one can better estimate the underlying support for the parties in the electorate.

To generate these estimates of "underlying support," I first examine shifts in the citizenry's overall evaluation of the parties, and second, and more important, I examine the changes in the public's evaluation of particular issues that make up these broader assessments. Thus, for instance, do citizens view the Democrats or GOP more favorably on such matters as the economy, foreign policy, or race relations? And has the Democratic (or Republican) lead on these issues expanded or contracted over the last ten years?

The Evidence

To measure change in the public's partisan evaluations, I rely on the National Election Studies' (NES) open-ended questions on the likes/dislikes of the two parties. The actual wording of these questions is:

—Is there anything in particular you like about the (Democratic or Republican) party? If yes, what is that? Anything else?

—Is there anything in particular you don't like about the (Democratic or Republican) party? If yes, what is that? Anything else?

Scholars have long recognized the merits of these questions. Stokes, Campbell, and Miller (1958), for instance, saw two major advantages in using these data. First, since respondents can answer questions in their own words, these questions tap the "diversity or relative frequency of ideas" (1958, 369). As a result, one can get a handle on concerns that are *salient* to the electorate (Kelley 1983; RePass 1971). This aspect of the data may be crucial, since scholars generally agree that an issue must be salient to citizens before it can alter their evaluations of the parties (Campbell et al. 1960; Carmines and Stimson 1981 and 1984; Sundquist 1983). Moreover, since the NES codes up to five comments per respondent for the four questions, one gets a good sense of the diversity of public opinion.[1] Of course, not all citizens make comments, but respondents, on average, make between three and four observations about the parties.

The second advantage of these data, noted by Stokes, Campbell, and Miller (1958), is that one can examine comparable data from each presidential election since 1952. The data currently span four decades, providing an excellent opportunity to study change in the electorate's partisan evaluations.[2]

Issues

The comments from the NES open-ended questions can be used to detect the influence of specific "issues" or concerns on the public's evaluation of the parties. The NES codes about 400 different responses. By grouping these comments to represent different political issues, one can estimate their influence on the public's evaluation of parties. The critical question becomes: How does one group these various comments? Stokes, Campbell, and Miller (1958) classified the comments according to the respondents' attitudes toward the candidates, group-related concerns, the parties as managers of government, domestic issues, and matters of foreign policy. Kelley (1983) expanded on work by Stokes and his collaborators, providing a more detailed breakdown of domestic and foreign policy issues. On the domestic front, for instance, he examined such concerns as the "election's economic impact," ideological stance, or medical care.

---

[1] In 1972, the CPS released only three comments per question for each respondent, rather than the normal five. In another version of the data, all five comments were included. I want to thank Michael Kagay for making this latter version available to me.

[2] Note that some scholars have serious doubts about the data generated by the like/dislike questions (see Lodge, McGraw, and Stroh 1989; Smith 1989). Other scholars see merit in these data (Geer 1988 and 1991; Kelley 1983; Knight 1985; RePass 1971; Wattenberg 1990). I prefer to avoid this debate here, and make only the claim that these data shed light on how the electorate judges the parties.

Building on these earlier works, I developed a coding scheme to examine the following sets of concerns: the economy; foreign policy; race relations; social-welfare matters; political philosophy; concerns about big business and labor; the management of government; and "contemporary" issues, which include references to crime, drugs, busing, abortion, gun control, and women's rights (see the appendix for details about the coding).[3] These issues, obviously, are only a subset of all possible concerns. But they cover many of the major disputes facing parties and should shed light on how the public evaluates the parties.

Estimating the Influence of Issues

To determine the impact of an issue on how the electorate evaluates the parties, we must first determine its salience and then estimate whether the public views one party more favorably than the other on that issue. Fortunately, Kelley (1983) provides a way to estimate the salience and the partisan bias of issues using the NES's open-ended questions. According to Kelley, salience is simply the proportion of respondents "who cite a given issue as something to like or dislike about the major parties . . ." (1983, 61).[4] Bias, on the other hand, is the proportion of citizens who favor one of the parties on a particular issue. To estimate bias, one first subtracts the pro-Republican comments from the pro-Democratic comments for that particular issue.[5] The result is a scale that can range from +10, a highly Democratic score, to −10, a highly Republican score. With this measure, one then divides the number of respondents who favor one party by all respondents who favor either party on that issue, yielding what Kelley calls bias (see Kelley 1983, 60–64).

With measures of salience and bias, one can then calculate the partisan influence of an issue. Here, I want to introduce a new concept,

---

[3] See Kessel 1984 for a different, but useful, coding scheme.

[4] One could quarrel with this measure of salience, since for an issue to be salient all an individual has to do is mention it once. Surely, other issues are likely to be important to citizens besides those mentioned in an open-ended format. The problem is that assessing the salience of an issue to the public is difficult. Closed-ended questions do not offer much of an alternative, since they also run into problems (Zaller and Feldman 1990). But, at least, the open-ended format does encourage respondents "to define [their] own issue space by naming issues that were salient to [them]" (RePass 1971, 391).

[5] There are two items to keep in mind when one adds and subtracts these comments. First, a "pro-Democratic" comment is either a statement about why a respondent likes the Democratic party *or* likes the Republican party. The reverse holds for "pro-Republican" comments. Second, in the calculation of these scales, I weigh each comment equally. For thoughtful defenses of the equal weighting assumption, see Kagay 1980 and Kelley 1983.

Net Pull, the relative ability of the two parties to attract supporters on a given issue. To determine the *pull* of an issue for a party, one multiplies the bias for a party by its salience. The result is the "proportion of . . . respondents who see a given issue as favoring a given . . . party" (Kelley 1983, 61). One then calculates the pull of any issue for *both* parties, subtracting the two proportions to yield the Net Pull. The specific formula is:

$$\begin{aligned}\text{Net Pull} = {} & (\text{salience of an issue} \times \text{Democratic bias on that issue}) \\ & - (\text{salience of an issue} \times \text{Republican bias on that issue})\end{aligned}$$

Estimates of pull for both parties are necessary, since highly salient issues may have a lot of pull for each party. For example, an issue that is salient to the public with a Democratic bias of 60% will still help the Republicans gain a more positive evaluation among the 40% who see the issue favoring the GOP. Therefore, looking only at the pull for one party provides a misleading picture of the overall influence of that issue on the electorate's partisan evaluations. Net Pull avoids this problem by providing direct insight into how an issue yielded change in the *net* partisan evaluations of the electorate. A useful way to think about this concept is that it estimates the lead one party has over the other on a particular issue.

One final methodological task involves estimating the public's *overall* evaluation of the two parties. Here I shall follow a simple procedure first adopted by Stokes, Campbell, and Miller (1958): namely, summing *all* the respondents' likes and dislikes of the two parties. This summation is used to generate three indicators of party support. First, "net Republican score" indicates the proportion of respondents who made more pro-Republican comments than pro-Democratic comments. "Net Democratic score," as one might guess, is just the reverse. "Neutrals," on the other hand, are those citizens who made no comments about either party or who made an equal number of pro-Democratic and pro-Republican comments.

The Findings

Table 1 reports the public's overall partisan assessments from 1952 to 1988. One obvious pattern in table 1 is that, from 1968 to 1980, the proportion of neutrals nearly doubled. This trend, which Wattenberg (1990) has carefully documented, suggests that both parties were further away from majority status in the 1980s than they were in the 1950s. My purpose, however, is to examine the shifts in the electorate's *relative* assessment of the parties, which is well summarized by the net Democratic edge. Here, the results suggest that the GOP has

Table 1. The Electorate's Overall Partisan Evaluations, 1952–88

	Net Democratic Score	Neutral Score	Net Republican Score	Net Democratic Edge	*N*
1952	44.4%	16.3%	39.3%	5.1%	1,799
1956	44.0	22.2	33.8	10.2	1,762
1960	44.7	21.6	33.6	11.1	1,932
1964	48.2	25.7	26.1	22.1	1,571
1968	37.1	23.8	39.1	−2.0	1,557
1972	35.3	35.9	28.9	6.4	1,372
1976	38.7	37.2	24.1	14.6	2,870
1980	33.9	41.1	25.0	8.9	1,570
1984	34.8	38.9	26.3	8.5	2,257
1988	36.1	34.8	29.1	7.0	2,040

NOTE: The proportions are the result of simply adding up the likes and dislikes of the parties. A net Republican score could range from −1 to −10. A net Democratic score could fall between 1 and 10. Individuals who either made no comments or whose comments cancelled each other out were given neutral scores. The net Democratic edge is derived by subtracting the net Republican score from the net Democratic score.

not made significant inroads into Democratic support. If one, for instance, looks at the net Democratic edge between 1952 and 1960, the average lead was about 9 percentage points. From 1980 to 1988, that lead was about the same—8 percentage points. This is not the only story one could tell, however. If one, instead, compares the net Democratic edge in 1964 to that of 1988, the GOP has made important gains. But it is unclear why the high point of Democratic support should be one's starting point. In contrast, one might compare 1968 to 1988. In that case, the Democrats have made sizable strides in the last twenty years. Using the low point of Democratic support also poses problems. These results, in sum, are open to a number of interpretations. But it is hard to escape one simple fact: the electorate continues to evaluate the Democrats more positively than the Republicans.

Table 2 provides a detailed breakdown of these evaluations, examining the partisan direction for eight issues across the last four decades. Of all the issues examined, concerns about big business–labor are clearly the most important to the electorate's assessment of parties. From 1952 to 1988, about 35% of the electorate, on average, mentioned concerns about labor or big business in their likes and dislikes about the parties. These attitudes have remained decidedly in favor of the Democrats during that time. The Democratic bias has consistently hovered at around 88%, generating net pulls of nearly 30%. No other issue

Table 2. Changes in the Electorate's Partisan Evaluations, 1952–88 (in percentages)

Issue	1952	1956	1960	1964	1968	1972	1976	1980	1984	1988
Government management										
Salience	25	9	11	14	15	10	14	14	12	19
Dem bias	7	18	34	42	29	45	34	19	33	36
Net pull	−22	−6	−4	−2	−6	−1	−5	−9	−5	−6
Foreign policy										
Salience	29	24	25	17	25	18	14	12	21	25
Dem bias	18	18	33	37	15	32	31	25	44	39
Net pull	−18	−15	−8	−4	−18	−7	−4	−6	−3	−6
Race										
Salience	5	6	8	15	9	3	3	3	2	2
Dem bias	63	60	72	54	68	70	80	98	82	61
Net pull	1	1	4	1	3	1	2	3	2	1
The economy										
Salience	60	32	34	21	30	20	26	19	24	31
Dem bias	56	62	54	58	41	64	66	37	46	39
Net pull	7	8	3	4	−5	6	9	−5	−2	−9
Big business–labor										
Salience	39	41	37	36	32	35	34	34	31	35
Dem bias	88	88	85	87	87	89	88	88	89	88
Net pull	30	31	36	26	24	27	26	25	24	27
Philosophy										
Salience	23	15	19	23	25	18	19	20	24	26
Dem bias	42	45	40	28	32	33	32	36	36	41
Net pull	−4	−2	−4	−11	−9	−6	−7	−5	−7	−5
Social welfare										
Salience	17	18	17	17	16	12	11	12	15	23
Dem bias	86	81	70	77	64	47	50	42	57	69
Net pull	12	11	7	9	5	−1	0	−2	2	9
"Contemporary" issues										
Salience	—	—	—	—	—	4	3	2	6	8
Dem bias	—	—	—	—	—	44	36	57	44	34
Net pull	—	—	—	—	—	−1	−1	1	−1	−3

NOTE: *Salience* is the proportion of respondents who made at least one comment about the issue. *Dem bias* is the proportion of respondents who see an issue favoring the Democrats. The Republican bias equals 100% minus the Democratic bias. Neutrals are excluded from this calculation. *Net pull* is the proportion of respondents "pulled" by the issue toward a particular party. A negative sign means the pull is toward the Republican party. Pull is arrived at by multiplying salience by bias for the party.

studied here has had such a powerful effect on the public's assessment of the parties. Interestingly, one foundation of the New Deal was that the Democrats were the party of the working man and woman, and that image appears to remain a centerpiece of the citizens' partisan evaluations—even 50 years after FDR set the New Deal in motion.

While the Democrats benefit greatly from that issue, other concerns have consistently favored the GOP. For instance, comments about foreign policy, political philosophy, and how the government is run have all exhibited a pro-Republican edge in each of the last ten presidential elections. The degree of support for the GOP varies from election to election, but the net pulls show a consistent GOP advantage. While the Republicans have positive ratings on these three issues, their lead has *not* increased in recent years. If one compares, for instance, the net pulls for these three issues in 1988 to the same measures in 1980 (or 1976), there is very little change.

The influence of concerns about social welfare, on the other hand, sends mixed signals. From 1972 to 1980, the GOP experienced sizable gains in the public's evaluation on this issue. During these three elections, the Democratic bias averaged about 46%. This shift was quite a reversal of fortunes: between 1952 and 1964 the Democratic bias was about 79%. But in the last two elections, the Democrats have reestablished an edge on these matters. In 1988, for instance, the net pull for the Democrats was about the same as it was in 1960 and 1964. According to these data, there does not appear to be a rejection of the Democrats' position on social-welfare concerns by the public in 1984 or 1988—one of the commonly cited cornerstones of an emerging Republican majority.

One potentially important change in favor of the GOP does, however, emerge on the economic front. Since 1980, the public has begun to view the GOP more favorably on matters pertaining to the economy. In 1976, for instance, the Democratic bias was 66%. By 1988, it had slipped to 39%, yielding a net pull of 9% for the Republicans. It appears that Carter's economic difficulties combined with Reagan's economic successes have led the public to think of the Republicans as the "party of prosperity." And since economic issues are often salient when citizens evaluate the parties, this change bodes well for the GOP.

The two remaining concerns, "contemporary" and racial issues, have simply not been salient enough to generate significant change in the public's evaluation of the parties. The low salience of "contemporary" issues has led to net pulls of just −1% in three of the last five elections. Racial issues follow a similar pattern. Between 1972 and 1988, only about 2% to 3% of the electorate made any references to racial concerns. With such low salience, these issues have been unable to recast the electorate's partisan evaluations. The only exception was 1964, when racial matters were salient to 15% of the public. But in that year the Democratic bias was just 54%, leading to a small overall impact for those concerns.[6]

---

[6] One may find this last set of findings surprising, given that race is generally viewed as a central issue in U.S. politics over the last 40 years (see Carmines and Stimson

The Newest Cohort

Even though tables 1 and 2 indicate that the Democrats continue to garner more favorable evaluations than the Republicans, perhaps the youngest members of the electorate have developed more positive evaluations of the GOP than of the Democrats. Because of generational replacement, these changes among the youth could fuel a future GOP majority, as Norpoth (1987) and Norpoth and Kagay (1989) suggest. To test this possibility, I examined the partisan evaluations of the 18–29-year-olds for the last three presidential elections.

The data in table 3, however, lend little support to the interpretation that the GOP has built a favorable impression among the youth of the 1980s. In each of the last three presidential elections, these citizens have favored the Democratic party. The Democratic edge ranged from 2 percentage points in 1980 to 12 percentage points in 1984. These young citizens do, however, tend to be more neutral in their assessment of parties than the entire public, suggesting that neither party has captured their imaginations. But even though these 18–29-year-olds have less to say about the parties, the partisan content of their responses is similar to that of the entire electorate (see table 4). Comments about labor and big business, for example, strongly favor the Democrats. For social welfare concerns, the Democrats maintained a small edge in net pull in 1984 and 1988, again indicating that, among these newest voters, there has not been a widespread rejection of the Democrats' position on social spending. On economic issues, the GOP holds a lead comparable to its lead among the entire electorate. All in all, it would be difficult to argue that the newest members of the electorate will provide a foothold for the GOP as it strives to become the majority party.

Conclusion

These results show that the Democratic party, despite recent defeats in presidential elections, remains the favored party. The Democratic edge in the public's overall partisan evaluations is about the same in

---

1989). These data may underestimate the influence of racial issues, especially since respondents may be hesitant to mention these concerns for fear of saying something socially unacceptable. This argument surely has merit, but there may be more to the story. In 1964, for instance, the salience of racial concerns rose to 15%, suggesting that a good number of citizens were willing to mention this matter in some situations. In addition, the proportion of citizens citing racial concerns has *dropped* since 1964, suggesting that racial issues are less salient to the public's partisan evaluations in the 1980s than they were in the 1960s.

Table 3. Changes in Partisan Evaluations of 18–29-Year-Olds, 1980–88 (in percentages)

	Net Democratic Score	Neutral Score	Net Republican Score	Net Democratic Edge	*N*
1980	25	52	23	2	422
1984	33	46	21	12	548
1988	29	46	25	4	450

NOTE: See table 1 for description of measures.

the 1980s as it was in the 1950s. In addition, the electorate's relative evaluation of the Democrats and Republicans on seven of the eight issues studied here have not changed greatly over the last 20 years. The Democrats continue to benefit greatly from concerns associated with big business and labor and also remain ahead on social-welfare matters. The only significant change occurs for economic issues, where the GOP have built a lead over the Democrats in recent years. Finally, the youth of the nation do not appear to be pro-Republican in their evaluations of the parties, suggesting that generational change may not lead to a GOP majority. These data, in sum, suggest that *continuity*, rather than change, is an important theme underlying the public's relative evaluations of the parties from 1952 to 1988.

This analysis also has more general lessons for the study of partisan change. Specifically, the idea that issues pull and tug on citizens' overall evaluations of parties highlight the many obstacles that face the minority party's climb to the top. First, table 2 shows that few issues have had great impact on the overall partisan evaluations of the electorate. Except for concerns about big business and labor, no issue has had an average net pull of more than 10% during the 36-year period. It makes sense that most issues will be salient to only a portion of the population, especially given the diversity of the U.S. electorate. But because of this diversity, most issues simply lack the power to reshape the partisan moorings of the entire electorate.

A second obstacle is that many issues confront the public at any one time—each tugging and pulling on the citizens' overall evaluations of the parties. The economic growth during the Reagan administration, for instance, has surely given many young voters reason to support the Republicans. But simultaneously, other positions of the GOP, such as those on social-welfare issues, may be retarding partisan change toward them. Since there is this pull-and-tug effect and few issues have widespread salience, it becomes clearer why it has been over 50 years

Table 4. Changes in Partisan Issue Evaluations of 18–29-Year-Olds, 1980–88 (in percentages)

Issue	1980	1984	1988
Government management			
Salience	12	10	14
Dem bias	19	43	40
Net pull	−7	−1	−3
Foreign policy			
Salience	13	24	27
Dem bias	27	53	36
Net pull	−6	1	−8
Race			
Salience	2	2	1
Dem bias	100	89	83
Net pull	2	2	1
The economy			
Salience	17	20	28
Dem bias	14	42	35
Net pull	−12	−3	−8
Big business–labor			
Salience	20	24	24
Dem bias	84	87	86
Net pull	14	18	17
Social welfare			
Salience	8	11	16
Dem bias	39	73	78
Net pull	−2	5	9
Philosophy			
Salience	17	20	20
Dem bias	46	54	57
Net pull	−1	2	3
"Contemporary" issues			
Salience	2	7	7
Dem bias	40	32	50
Net pull	−1	−3	0

NOTE: See table 2 for description of measures.

since a new majority party emerged. The forces associated with low-salience issues can only chip away at the partisan evaluations of the electorate.

This argument strongly suggests that permanent, large-scale shifts in the overall partisan evaluations of the public may be impossible without *both* a highly salient issue that dominates the concerns of most citizens and an issue that is perceived by the public as heavily favoring one party—just as Sundquist (1983) and others have argued. The Great Depression and FDR's response to it apparently brought such a shift. Without a shock of similar magnitude, partisan evaluations will change only slowly and modestly, as Key (1959) observed over 30 years ago. Given the logic of this analysis, the pro-Democratic evaluations may well continue. And at least at this point, Reagan's legacy, although important, does not appear to be strong enough to have significantly altered the overall partisan evaluations of the U.S. public, suggesting that the Democrats have reason for optimism as they enter the decade of the 1990s.

Appendix 1

Open-ended comments are recoded into eight different groups of issues. A list of the specific codes for each issue and a sample of some of the kinds of comments in the categories is given below. For a detailed account of Kelley's coding scheme for the 1964 and 1972 elections, see Kelley 1983 (239–62). No measure of coding reliability is available, since only a single coder (the author) was used.

I have only presented the coding scheme for 1984. From 1972 to 1988, the NES's coding arrangement has been nearly identical. Prior to 1972, however, the codes varied a bit from year to year. The actual content of the responses did not change, but the numeric code assigned to the responses did shift. Exact codes for the years before 1972 can be obtained by writing to the author.

The Economy: 605, 606, 901, 926 to 941, 1007 to 1009
- The party would spend less/more
- The party would lower/raise taxes
- The party supports a government program to create jobs
- Inflation would be better/worse under the party

Social-Welfare Policies: 905–15, 920–25, 942–54, 1001, 1003, 1215–16, 1219–22
- The party favors welfare/poverty programs
- The party favors expansion of social security/pensions
- The party is against aid to education
- The party is against national health insurance

Government Management: 601–4, 607–20, 697
- The party runs a good/efficient administration
- The party runs honest government

—The party supports reforming the bureaucracy
—The party works well with Congress
Foreign Policy: 1101–97
—The party favors a strong military
—The party handles trouble spots well
—The party handles Russia well
—The party favors a nuclear freeze
Big Business–Labor: 952–58, 1201, 1202, 1205–10, 1233, 1234
—The party favors privileged people
—The party supports the common man/working people
—The party supports unions
—The party helps the poor
Race: 946–48, 1217–18
—The party favors civil rights
—The party is opposed to integration
—The party supports voting rights
—The party favors blacks
Political Philosophy: 162–65, 805–28, 833–34, 837–38, 902, 904
—The party is against social change
—The party favors the work ethic
—The party is opposed to a planned economy
—The party is too liberal
Contemporary Issues: 968–86, 988–96, 1223–26
—The party supports law and order
—The party is too permissive
—The party favors a woman's right to an abortion
—The party is opposed to gun control

References

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

Carmines, Edward G., and James A. Stimson. 1981. "Issue Evolution, Population Replacement and Normal Partisan Change." *American Political Science Review* 75:107–18.

Carmines, Edward G., and James A. Stimson. 1984. "The Dynamics of Issue Evolution: The United States." In *Electoral Change in Advanced Industrial Democracies*, ed. Russell Dalton, Scott Flanigan, and Paul Allen Beck, 134–58. Princeton: Princeton University Press.

Carmines, Edward G., and James A. Stimson. 1989. *Issue Evolution*. Princeton: Princeton University Press.

Cavanagh, Thomas, and James Sundquist. 1985. "The New Two-Party System." In *The New Direction of American Politics*, ed. John E. Chubb and Paul E. Peterson, 33–68. Washington, DC: Brookings Institution.

Geer, John G. 1988. "What Do Open-ended Questions Measure?" *Public Opinion Quarterly* 52:365–71.

Geer, John G. 1991. "Do Open-ended Questions Measure 'Salient' Issues?" *Public Opinion Quarterly*. In press.

Kagay, Michael. 1980. "On Two Models of Voter Decision Making: Stokes's Six Components and the Kelley-Mirer Rule." Princeton University. Photocopy.

Kelley, Stanley, Jr. 1983. *Interpreting Elections*. Princeton: Princeton University Press.

Kessel, John H. 1984. *Presidential Parties*. Homewood, IL: Dorsey Press.

Key, V. O. 1959. "Secular Realignment and the Party System." *Journal of Politics* 21:198–210.

Knight, Kathleen. 1985. "Ideology in the 1980 Election: Ideological Sophistication Does Matter." *Journal of Politics* 47:828–53.

Lodge, Milton, Kathleen M. McGraw, and Patrick Stroh. 1989. "An Impression-driven Model of Candidate Evaluation." *American Political Science Review* 83:399–420.

Norpoth, Helmut. 1987. "Under Way and Here to Stay: Party Realignment in the 1980s?" *Public Opinion Quarterly* 51:376–91.

Norpoth, Helmut, and Michael Kagay. 1989. "Another Eight Years of Republican Rule and Still No Partisan Realignment?" Paper presented at the annual meeting of the American Political Science Association, Atlanta, GA.

Petrocik, John R. 1987. "Realignment: The South, New Party Coalitions, and the Election of 1984 and 1986." In *Where's The Party?* 21:33–59. Washington, DC: Center for National Policy.

RePass, David E. 1971. "Issue Salience and Party Choice." *American Political Science Review* 65:389–400.

Shanks, Merrill J., and Warren E. Miller. 1989. "Alternative Interpretations of the 1988 Election." Paper presented at the annual meeting of the American Political Science Association, Atlanta, GA.

Smith, Eric R. A. N. 1989. *The Unchanging American Voter*. Berkeley: University of California Press.

Stokes, Donald, Angus Campbell, and Warren E. Miller. 1958. "Components of Electoral Decision." *American Political Science Review* 52:367–87.

Sundquist, James L. 1983. *Dynamics of the Party System*. Washington, DC: Brookings Institution.

Wattenberg, Martin P. 1987. "The Hollow Realignment: Partisan Change in a Candidate-centered Era." *Public Opinion Quarterly* 51:58–74.

Wattenberg, Martin P. 1990. *The Decline of American Political Parties*. 3d ed. Cambridge, MA: Harvard University Press.

Zaller, John, and Stanley Feldman. 1990. "Answering Questions vs. Revealing Preferences: A Simple Theory of the Survey Response." Paper presented at the annual meeting of the Political Methodology Society.

THE SHORT-TERM EFFECT OF ELECTION POLLS ON FOREIGN EXCHANGE RATES: THE 1988 CANADIAN FEDERAL ELECTION

GARY A. MAUSER and
CRAIG FITZSIMMONS

Abstract Polls do not simply measure public opinion; they also provide vital information that the public can use to form opinions and to make decisions. Using multiple regression analysis, the authors found that published polls had a powerful impact on the value of the Canadian dollar during the 1988 Canadian federal election. This appears to have been due to the unprecedented importance of economic issues in the federal campaign, the distinctive positions taken by the major political parties with respect to the U.S.-Canadian Free Trade Agreement, and to the proliferation of "horse-race" polls in the media. These results have implications for understanding the behavior of other elite groups, such as financial contributors and volunteer campaign workers, who may also use published horse-race polls to handicap election outcomes.

Polls do not simply measure public opinion; they also provide vital information that the public can use to form opinions and to make decisions. This is most evident during elections, when poll results are an increasingly important part of media coverage, but it can also be seen between elections when pressure groups use poll results to support their demands for public funds. Researchers have argued that published polls can influence a variety of phenomena, the voting intentions of individual voters, for example, or the morale of campaign

GARY A. MAUSER is Professor of Business Administration at Simon Fraser University. CRAIG FITZSIMMONS is a Media Researcher at the Fraser Institute in Vancouver, B.C. The authors would like to thank Dennis Maki and John Richards, both of whom are at Simon Fraser University, and Isabella Horry at the Fraser Institute for their comments on an earlier draft of this paper. Please direct all communications to Gary Mauser, Faculty of Business Administration, Simon Fraser University, Burnaby, B.C., Canada V5A 1S6.

Public Opinion Quarterly 55:232–240 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-02/$2.50

workers during elections. The purpose of this paper is to demonstrate how published opinion polls can influence money traders and, by extension, members of any other elite groups who would be motivated to read and act upon poll results, such as campaign contributors or volunteer campaign workers.

A money trader's decisions to buy or to sell a particular currency are based upon his or her view of the future movement of the exchange rate, so that virtually any newsworthy economic or political development can affect "spot" exchange rates (Levi 1983). Economists have had difficulty understanding the short-term causal factors. Some have even proposed a random-walk model (Backus 1984; Meese and Rogoff 1983). We contend that, during the 1988 Canadian federal election, the value of the Canadian dollar shifted in response to the rise and fall of support for the Progressive-Conservative (PC) party in published public opinion polls.

Polls and the Short-Run Foreign Exchange Market

In an age of communications, foreign currency traders have become information dealers (Grabbe 1986). Since a trader's decisions to buy or to sell are based upon his or her view of the future movement of the exchange rate, any new piece of information could influence trader expectations. Economists argue that the most important factor in understanding long-term movements in exchange rates is the relative inflation rates of the two countries whose currencies are being traded. That is, if the rate of inflation in Canada is higher than that in the United States, in the long run, we would expect that the value of the Canadian dollar relative to the U.S. dollar—its exchange rate—should decline. The best index of inflationary pressure on the Canadian dollar is the spread between Canadian and U.S. three-month treasury bills (Lipsey, Purvis, and Steiner 1988). The spread between U.S. and Canadian treasury bills reflects the countries' different interest rates. As a matter of monetary policy, Canadian treasury bills pay higher rates than U.S. treasury bills in order to entice investors.

However, short-run forces often prevail over long-run forces. In 1988, the Canadian federal election was dominated by the U.S.-Canadian Free Trade Agreement. Prime Minister Brian Mulroney called the Canadian federal election on October 1, 1988, and the election was held on November 21, 1988. The Progressive-Conservative party strongly backed the agreement; indeed, they had just recently negotiated it. Due to the Free Trade Agreement, the business community enthusiastically rallied behind the Conservatives.[1] A well-financed business lobby group was set up prior to the campaign to promote the Free Trade Agreement and, indirectly, the Conservatives. This group was called the Canadian Alliance for Trade and Job Opportunities (CATJO) and comprised the Canadian Federation of Independent Business, the Consumers' Association of Canada, the Business Council on National Issues, the Canadian Manufacturer's Association, the Canadian Chamber of Commerce, and the Canadian Exporters' Association. CATJO spent over $4 million (Canadian) during and before the election to promote the Free Trade Agreement. Both opposition parties, the Liberal party and the smaller New Democratic party, vehemently opposed the Free Trade Agreement and pledged to dissolve the agreement if they were to form the government. The debate about the Free Trade Agreement quickly came to dominate the campaign. Such unanimity about what is the major issue is rare in Canadian politics.

The Economic Council of Canada predicted that free trade would improve real wages, increase employment, stimulate business investment and productivity, lower prices, reduce government budget deficits, and strengthen the Canadian dollar (Magun, Rao, and Lodh 1987). Economic analysis suggested that freer trade between Canada and the United States would lead to greater investment from third countries, including Japan and the European nations (Rugman 1988). In such an economic environment, the financial communities in both the United States and in Canada would be expected to associate the passage of the Free Trade Agreement with a stronger Canadian economy and be encouraged to purchase Canadian dollars if they expected free trade would become a reality. However, because passage of the Free Trade Agreement was dependent upon the Progressive-Conservatives winning a majority in Parliament, we hypothesized that, during the election campaign, money traders would use public opinion polls as their primary information source in determining whether or not to purchase Canadian currency.

Methodology

To test the hypothesis that public opinion polls influenced the foreign exchange rate of the Canadian dollar during the Canadian federal election campaign, multiple regression equations were defined using the value of the Canadian dollar (in U.S. dollars) as the dependent variable. Specifically, the dependent variable was the average exchange rate of the Canadian dollar for up to three days following the publication of each poll. Fewer days were included when the subsequent poll was published within three days.

---

[1] The business community historically has been divided between the Liberals and the Conservatives, but in 1988, due to John Turner's strong opposition to the Free Trade Agreement, business support for the Liberals evaporated (Frizzell, Pammett, and Westell 1989, 69).

The independent variable of key importance was the popular support for the Progressive-Conservatives (PCs) as measured by the polls published in the major Canadian news media over the 14 weeks leading up to the 1988 election. During this period, beginning more than one month before the official start of the campaign, there were 27 horserace polls published between August 11 and November 19, 1988. Support for the PCs was selected since it was the only party that supported the Free Trade Agreement. Only reputable, nationwide polls were included that were repeated throughout the campaign: Environics/*Globe and Mail* (4), Insight Canada/CTV Network (6), Gallup Canada (9), Canadian Facts/CBC Network (2), and Angus Reid/*Southam News* (6). The regression model contains 20 points because a few polls were averaged together when more than one poll was reported on the same day or over a weekend.

To evaluate the inflationary pressure on the Canadian dollar, a second independent variable was included in the regression model. Since the United States is Canada's largest trading partner, the spread (difference) between U.S. and Canadian treasury bills was included as a proxy for the relative health of the Canadian economy. In this study, the U.S. and Canadian three-month treasury bill rates were taken to be the rate fixed at noon on the day each of the 27 polls were published, or the nearest preceding trading day.[2] Thus, each of the independent variables was measured at least one day prior to the determination of the value of the Canadian dollar.

Results

As may be seen in table 1 and figure 1, the regression analysis demonstrates that there was a strong relationship between popular support for the Progressive-Conservatives and the value of the Canadian dollar across the period of the 1988 Canadian federal election campaign.

Support for the Progressive-Conservatives, as published in polling results, was the strongest factor, significant at less than .001 using a two-tailed t-test. The spread between Canadian and U.S. treasury bills,

---

[2] The three-month T-bill rates were drawn from the daily Report on Business section in the *Globe and Mail*.

Table 1. OLS Regression: Explaining Fluctuations in the Canadian Dollar during the 1988 Election

Independent Variable	Standardized Coefficient	Standard Error	t-ratio (17 df)
PC support	0.816*	0.035	5.318
T-bill spread (U.S.-Can.)	−0.070	0.810	−0.459
R^2	0.636		
F	14.841		
Signif. F	0.0001		

* $p < .001$.

the other independent variable, was not found to be significant. These results strongly support the hypothesis that Canadian money traders used the polls to handicap the results of the election and, therefore, to estimate the future value of the Canadian dollar.

This finding appears to be due to the exceptional clarity of what decision money traders should make based on the polls. Because the Conservatives were the only political party supporting the Free Trade Agreement, and given the unprecedented importance of the U.S.-Canadian Free Trade Agreement in the federal campaign, money traders could use the PC's standings in the polls to handicap the election outcome.

As may be seen in figure 1, the value of the Canadian dollar followed the growth in support for the Conservatives, during the first few months, and then it trailed the Tory popularity down immediately after the nationally televised debates between the party leaders on October 24 and 25. This decline reflected the "victory" in these debates of John Turner, the Liberal leader, over Brian Mulroney, the Progressive-Conservative leader. Equally important, the dollar began to rise again in early November after the resurgence of the Conservatives in the polls.

In order to determine the necessary conditions for this effect, a parallel analysis was conducted of the 1984 Canadian federal election. This election met many of the conditions hypothesized for this effect to occur, but not all. First, there were numerous (17) horse-race polls published in the months prior to the 1984 election campaign.[3] Second,

---

[3] The regression model included only 14 points because two polls were unavailable and two were averaged together.

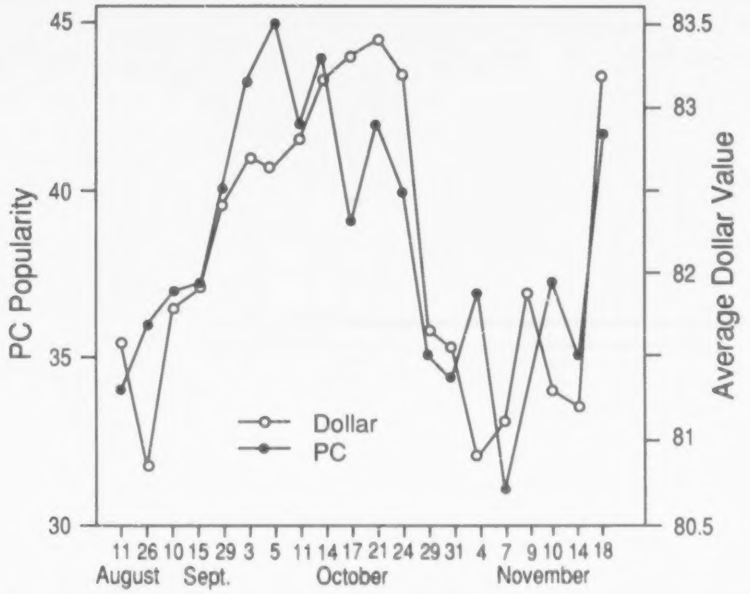


Figure 1. Average Canadian Dollar Value and Conservative Party Popularity in 1988

the business community favored the Progressive-Conservatives in 1984, although not with the same unanimity as in 1988. In 1984, the Liberals had been the party of government for many years and had attracted the support of many business leaders. But economic issues were not as important in the 1984 campaign as they were in the 1988 campaign. The 1984 campaign focused on political leadership: both the Progressive-Conservatives and the Liberals had recently nominated new leaders. Consequently, it was not easy for money traders to know which party would be best for the Canadian economy or for the dollar.

As expected, the analysis of the 1984 election campaign did not find a significant relationship between published horse-race polls and the value of the Canadian dollar (see table 2). Compared with the 1988 analysis, the coefficient for PC support in 1984 is significantly smaller at the 0.05 level ($t > 1.770$, $df = 32$) using a one-tailed test.[4] Indeed, during the 1984 campaign, the "spot" rate was driven by the spread between U.S. and Canadian treasury bills. In short, money traders were divided on which party would be best for the country's economy, as was the rest of the electorate, so they could not use political polls to handicap the expected value of the Canadian dollar. While it is

---

[4] A one-tailed test was used because the hypothesis is directional. Given the small sample sizes, it is reasonable to set .05 as the required level of confidence.

Table 2. OLS Regression: Explaining Fluctuations in the Canadian Dollar in the 1984 Canadian Election

Independent Variable	Standardized Coefficient	Standard Error	*T*-ratio (11 *df*)
PC support	0.046	0.017	0.425
T-bill spread (U.S.-Can.)	−0.924*	0.150	−8.425
R^2	0.880		
F	40.313		
Signif. *F*	0.0001		

* $p < .001$.

theoretically difficult to understand what drives the "spot" rate in the short term, during the 1984 campaign the Canadian dollar was at an all-time low and was particularly sensitive to variations in differences between U.S. and Canadian interest rates. As may be seen in figure 2, no relationship is apparent between the popularity of the Progressive-Conservatives and the value of the Canadian dollar.

Conclusions

The results of this study support the hypothesis that published polls had a powerful impact on the value of the Canadian dollar during the 1988 Canadian federal election. This appears to have been due to the unprecedented importance of the U.S.-Canadian Free Trade Agreement in the federal campaign, the exceptional clarity of what action to take as a result of the polling results, and to the proliferation of horse-race polls in journalistic accounts of the campaign. The degree of uncertainty about the outcome of the election during a campaign is another factor to consider. In 1984, the Conservatives held a commanding lead over the Liberals for much of the campaign, but in 1988 the lead shifted back and forth throughout the campaign, sometimes favoring the Liberals, sometimes the Conservatives. It is possible that this difference in relative certainty about the election outcome, coupled with the strikingly different positions taken on free trade, played a role in the thinking of the money traders.

These results have implications for understanding the behavior of other elite groups, such as financial contributors and volunteer campaign workers. Since volunteers and financial contributors are vital to almost every political campaign, demonstrating that polls influence

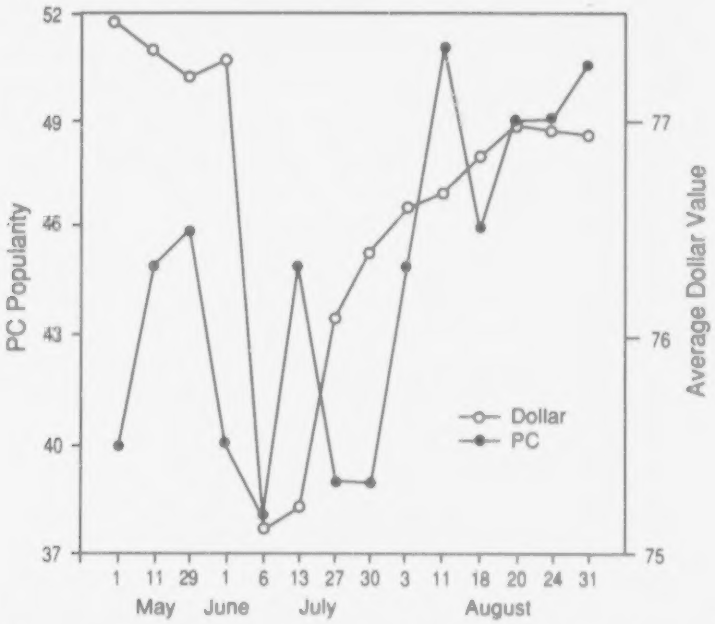


Figure 2. Average Canadian Dollar Value and Conservative Party Popularity in 1984

such groups would be an important step in showing that published polls could alter the outcome of elections. Clearly, financial contributors would be motivated to use published horse-race polls to handicap election outcomes in order to decide whom to support. Published polls may also influence campaign workers. Since campaign workers tend to be better educated and to be interested in politics, they would be expected to pay close attention to election news, including polls. It has been argued that volunteer campaign workers are likely to become demoralized if they see their candidate doing badly in the polls, and, therefore, they put less effort into the campaign (Bogart 1972). A parallel concern is that volunteers may become complacent if polls report their candidate too far ahead. Perhaps this can be restated more generally: the hypothesis is that the morale of campaign workers is driven, in part, by the competitiveness of the standing of their candidate in the polls. That is to say, volunteers are motivated to work harder when the polls show their candidate running neck-and-neck with the opposition, and that volunteers will tend to slack off when their candidate is either far behind or far ahead. Despite the seeming obviousness of these hypotheses, empirical support for them has been elusive.

Understanding the behavior of money traders would be expected to be a simpler problem than that of political activists, because of their widespread belief that the Canadian dollar would increase in value were the Conservatives to be reelected. Political activists, on the other hand, would be expected to have a more diverse set of reasons for being involved. Some might be involved for economic reasons and, hence, prefer to jump on the bandwagon of a winner, while others could be involved for ideological reasons and would be immune to polling results. However, it is possible that this problem may be less acute than previously thought. In the 1972 U.S. congressional elections, it was shown that economic motives dominated ideological reasons for contributing money to political candidates in that the bulk of funds came from contributors who preferred to support likely winners (Welch 1981). While Welch's study was based on election results rather than polling data, it suggests that it may be possible to demonstrate that campaign contributions vary temporally with a candidate's standing in the polls.

References

Backus, David. 1984. "Empirical Models of the Exchange Rate: Separating the Wheat from the Chaff." *Canadian Journal of Economics* 4:824–46.

Bogart, Leo. 1972. *Silent Politics: Polls and the Awareness of Public Opinion*. New York: Wiley-Interscience.

Frizzell, Alan, Jon H. Pammett, and Anthony Westell. 1989. *The Canadian General Election of 1988*. Ottawa: Carleton University Press.

Grabbe, J. Orlin. 1986. *International Financial Markets*. New York: Elsevier.

Levi, M. 1983. *International Finance*. New York: McGraw-Hill.

Lipsey, Richard G., Douglas D. Purvis, and Peter O. Steiner. 1988. *Economics*, 6th ed. New York: Harper and Row.

Magun, Sunder, Somershwar Rao, and Bimal Lodh. 1987. *Impact of Canada-United States Free Trade on the Canadian Economy*. Ottawa: The Economic Council of Canada.

Meese, Richard, and Kenneth Rogoff. 1983. "Empirical Exchange Rate Models of the Seventies: Are Any Fit to Survive?" *Journal of International Economics* 14:3–24.

Rugman, Alan. 1988. *Trade Liberalization and International Investment*. Ottawa: Economic Council of Canada.

Welch, W. P. 1981. "Money and Votes: A Simultaneous Equation Model." *Public Choice* 36:209–34.

SELF-INTEREST AND PUBLIC OPINION TOWARD SMOKING POLICIES: A REPLICATION AND EXTENSION

RICHARD D. DIXON,
ROGER C. LOWERY,
DIANE E. LEVY, and
KENNETH F. FERRARO

Abstract Numerous studies report that self-interest has a nonsignificant influence upon various political and social attitudes. In contrast, a recent article by Green and Gerken (1989) reports a California study showing that cigarette smokers are significantly more opposed than nonsmokers to public smoking restrictions and tobacco sales tax increases. The present article replicates and extends this analysis with data from two different states—Illinois (in 1984, *N* = 458) and North Carolina (in 1985, *N* = 488)—and using analysis techniques that differ from and expand upon the California study. Despite these methodological differences, self-interest is again shown to have a significant influence on opinions concerning public smoking restrictions and taxation, as well as on several additional issues relating to smoking and tobacco.

Introduction

Green and Gerken (1989) called attention to 22 studies indicating that self-interest did not have a significant influence upon political and social attitudes. Those studies revealed that directly affected individuals were neither more nor less likely than others to favor or oppose such things as jobs programs, government health insurance, aid to education, preferential hiring, the ERA, affirmative action for women, or raising the legal drinking age.

RICHARD D. DIXON is Professor of Sociology at the University of North Carolina at Wilmington; ROGER C. LOWERY is Professor of Political Science at the University of North Carolina at Wilmington; DIANE E. LEVY is Professor of Sociology at the University of North Carolina at Wilmington; and KENNETH F. FERRARO is Associate Professor of Sociology at Purdue University.

Public Opinion Quarterly 55:241–254 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-02/$2.50

In contrast, Green and Gerken's own analysis of data collected in two statewide telephone surveys of Californians concluded that self-interest did have a significant influence on attitudes toward public smoking restrictions and tobacco sales taxes. Data from a 1984 California Poll[1] revealed that support for public smoking restrictions was: (1) strongly and inversely related to cigarette consumption (''former'' and ''never'' smokers responded similarly), and (2) strongly and directly related to the extent that cigarette smoke bothered respondents. An increase in the polarization of respondents' attitudes resulted when the two independent variables were combined. A 1987 California Poll, concerning support or opposition to increases in tobacco sales taxes, if taxes had to be raised, indicated that nonsmokers were considerably more likely than smokers to support cigarette tax increases, with little regard for the level of increase. Conversely, smokers were evenly split concerning a five-cent-per-pack cigarette tax increase, but were increasingly unreceptive to the idea of larger tax increases. The relationship between respondent income differences and support for higher tobacco sales taxes was trivial for nonsmokers, but strong for smokers.

Green and Gerken attempted to explain why attitudes concerning public smoking restrictions and taxation constitute an exception to the lack of relationship between self-interest and policy attitudes reported in other research. Calling attention to several studies indicating that self-interest has a greater influence upon policy attitudes when the costs and benefits of a given proposal are unambiguous and fixed prominently in the minds of respondents, they argued that public smoking restrictions and tobacco sales taxes fall into this category. That is, public smoking restrictions and tobacco sales taxes involve unusually clear and salient stakes for respondents depending on whether or not they are smokers and whether or not they are bothered by smoke.

Purpose

Our study is a replication and extension of the Green and Gerken inquiry, using survey data from two states—Illinois and North Carolina.[2] We also examine the influence of smoking status and consequences on respondent views toward public smoking restrictions and tobacco sales taxes. However, our study goes further by including

---

1. Respondents were asked if smoking should be unrestricted, restricted to special areas, or banned in buses, restaurants, planes, workplaces, and hotels.

2. North Carolina is especially relevant to examining self-interest linkages to smoking policy preferences because it is the leading tobacco growing and producing state.

several additional dependent variables about tobacco relating to ideological and policy issues. Our study also includes numerous additional independent variables concerning individual and group-level self-interest (e.g., profiting from tobacco, parents' and best friends' smoking status, etc.), and a larger complement of sociodemographic variables such as education and state of residence.

Data Collection

In February, 1985, Dixon, Levy, and Lowery conducted a statewide random-digit-dial telephone poll of respondents 18 years old and older in North Carolina ($N = 488$). In October and November, 1984, a companion statewide poll was done by Ferraro in Illinois ($N = 458$). Respondent contacts occurred during evening hours, each interview lasting approximately ten minutes. Both survey instruments included numerous questions about tobacco use, public and private restrictions on smoking, the tobacco industry, and the health effects of smoking. The response rate among eligible individuals was 73% in Illinois and 71% in North Carolina, calculated by determining the proportion of the eligible individuals who provided enough information to be retained as usable cases.

Methods

Twelve dependent variables and fourteen independent variables are examined for this report. Each of the dependent variables is identified and briefly described in the appendix. Dependent variables are grouped into three categories: (1) direct interest issues, (2) other policy/program issues, and (3) ethical/legal issues. Regarding the first set of direct interest issue variables, a "Restrictions on Smoking in Public" scale similar to the Green and Gerken design was composed by summing responses concerning favorability (+1) or opposition (−1) to the restriction or prohibition of smoking in each of six settings. The −6 to +6 scale that was generated was converted to a −1 to +1 scale, a range of values similar to the other dependent variables. The one remaining direct interest issue dependent variable measures attitudes toward changing tobacco sales taxes. Three other policy/program issues concern federal policies influencing the production or marketing of tobacco. Finally, the authors extracted two additional ethical/legal issues from a much longer list of other opinion statements about smoking and tobacco because they appeared to have strong public policy implications.[3]

The appendix also identifies and briefly describes the fourteen independent variables considered in this report. Three of these are current direct interest variables in that they concern the current involvement of respondents/close relatives with smoking and tobacco products. Four other tobacco-related independent variables concern respondent histories, influences, state of residence, and problems relating to smoking and tobacco. Finally, there are seven sociodemographic variables that are routinely employed in social scientific research.

Results

SMOKING ISSUE OPINIONS

Table 1 displays the basic findings from this inquiry. The table shows percentages of respondents taking anti-tobacco positions on each of the 6 "Restrictions on Smoking in Public" scale items and on each of the other 6 smoking issue opinions dependent variables. Also included in the table is a decomposition of anti-tobacco position percentages on these 12 issues across categories of 7 measures of self-interest independent variables.

The first two independent variables are our measures of cigarette consumption and the extent to which cigarette smoke bothers respondents. Concerning "Smoke Status" and "Smoke Bother," the present data yield the same results as those found by Green and Gerken: both measures of self-interest clearly and significantly relate to attitudes toward smoking restrictions and tobacco sales taxes. "Smoke Status" and "Smoke Bother" also relate significantly to all five of the additional smoking policy issues that we examined. Moreover, as in the Green and Gerken findings, current "never" and "former" smokers are alike in their opinions on smoking and tobacco issues.

Extending the analysis to include five additional self-interest variables provides further evidence of the importance of self-interest as a factor influencing opinions on smoking policy issues. Inspection of the table reveals that opposition to tobacco is consistently higher among respondents who do not profit from tobacco, whose parents did not smoke during respondents' youth, whose families have had serious health problems due to smoking, and who have fewer best friends

---

[3] Each dependent variable was scored so that negative values indicate a pro-tobacco position, and positive values indicate an anti-tobacco position. Thus, the higher the value, the more anti-tobacco the position taken.

Table 1. Percentage of Respondents Demonstrating Anti-tobacco Positions, 1984–85

		Measures of Self-Interest																	
		Current Direct Interest Variables								Other Tobacco-Related Variables									
		Smoke Status			Smoke Bother[a]			Profit from Smoke		Smoke Health Problem		Best Friends Smoke				Parent Smoked		State	
Smoking Issue Opinions	Total	Never	Former	Smoker	Never	Sometimes	Often/Always	Yes	No	Yes	No	0	1	2	3	Yes	No	Ill.	N.C.
Direct interest issues																			
Restrict smoking (% favor)																			
In restaurants	80	88	88	62***	59	84	91***	64	82***	90	77***	91	83	80	66***	80	82	86	76***
In gov't buildings	69	75	76	53***	53	68	86***	45	71***	77	66**	79	72	63	58***	68	71	74	64**
In office/building workplaces	72	80	81	52***	54	74	84***	56	73***	78	69**	80	75	68	62***	70	76†	75	69*
Prohibit Smoking (% Favor)																			
In grocery stores	63	70	68	49***	47	62	81***	47	65***	67	62	70	65	61	54**	61	69*	76	51***
In some hotel/motel rooms	77	80	82	67***	67	77	85***	73	77	80	75	82	80	79	66***	74	81*	80	73*
On airline flights <2 hrs.	69	76	78	50***	55	68	83***	60	70†	74	67*	76	76	63	58***	66	76**	73	65**
Sales taxes on tobacco products (% Increase)	33	44	43	10***	15	33	55***	26	34	47	28***	45	38	30	21***	34	34	35	32
Other policy/program issues																			
Federal programs to aid tobacco farmers to switch (% Increase)	73	79	76	60***	65	72	83***	68	74	82	70***	81	80	71	59***	74	70	72	74
Federal tobacco crop support programs (% Decrease)	34	34	41	26**	29	33	40†	19	35***	42	30**	39	33	33	28*	34	31†	40	28***
Complete ban on cigarette advertising[b] (% Agree)	45	51	59	28***	19	43	78***	33	48**	54	42**	55	48	44	35**	41	54**	—	45
Ethical/legal issues																			
Right to smoke in public places (% Disagree)	26	33	31	11***	12	25	43***	16	27*	33	24**	38	27	20	17***	23	34***	28	24
Morally wrong to profit from tobacco business (% Agree)	23	30	24	12***	15	17	48***	15	24†	28	22†	33	18	20	19***	20	29**	24	22

NOTE: Analysis *N*s ranged from 834 to 921 cases when using data for both states; chi-square test results shown.

[a] For percentages shown, Smoke Bother scale (0–9) grouped as follows: Never = 0, Sometimes = 1–4, Often/Always = 5–9.

[b] North Carolina data only. *N*s ranged from 453 to 462.

† $p \leq .10$. * $p \leq .05$. ** $p \leq .01$ *** $p \leq .001$.

who smoke. Illinois residents also are more inclined to restrict public smoking and decrease federal crop support programs than are residents of "the tobacco state."

Significant differences aside, table 1 reveals that two smoking policy issues elicit a substantial degree of consensus in that majorities of respondents—frequently exceeding 60%—favor smoking restrictions and federal programs to help tobacco farmers switch. This is not the case for any of the other five issues, however, where—significant differences aside—less than a majority of the respondents take the antitobacco position.

BIVARIATE ASSOCIATIONS

Bivariate analysis of each dependent variable[4] was conducted by using, depending on the level of measurement, either one-way analysis of variance (eta), τ_b, or product-moment correlations (*r*).[5] These procedures revealed that the majority (65) of the 98 (7 × 14) bivariate associations involving these variables are significant at the .10 level or higher, as summarized in table 2.[6] With one exception, all dependent variables are significantly associated with at least 9 of the 14 independent variables. However, the influence of the independent variables demonstrates greater variation. Whereas 4 variables demonstrate significant associations with all 7 dependent variables, there also are 3 independent variables that are significantly associated with fewer than half of the dependent variables.

Table 2 also provides readers with an indication of which independent variables were then selected for inclusion in subsequent multivariate analysis of each dependent variable. Specifically, on theoretical grounds, all measures of self-interest were used irrespective of whether their bivariate associations were significant, but measures of sociodemographic status were used only in instances where their bivariate associations were found to be significant.

---

[4] For this and the subsequent multivariate analyses, the six public smoking restriction/prohibition items are combined into the "Restrictions on Smoking in Public" scale described earlier.

[5] Analyses involving smoking status initially were conducted using a five-category form of that variable: nonsmokers, former smokers, light smokers, medium smokers, and heavy smokers. However, the first two categories of individuals generally did not differ from each other, and the latter three categories also did not differ from each other. Therefore, smoking status responses were dichotomized into smokers and nonsmokers.

[6] We use the .10 level of significance as a determining basis for subsequent multivariate consideration of any independent variable—rather than the more conventionally accepted .05 level—to avoid premature deletion of variables that appeared to have a substantial impact upon one or more dependent variables. In fact, as table 2 reveals, most significant relationships were at the .01 level or higher.

Table 2. Smoking Issue Opinions by Measures of Self-Interest and Sociodemographic Status: Bivariate Associations, 1984–85

	Measures of Self-Interest							Measures of Sociodemographic Status						
	Current Direct Interest Variables			Other Tobacco-related Variables										
Smoking Issue Opinions	Smoke Status (Eta)	Smoke Bother (r)	Profit from Smoke (Eta)	Smoke Health Problem (Eta)	Best Friends Smoke (r)	Parent Smoked (Eta)	State (Eta)	Age (r)	Educ. (Tau$_b$)	Gender (Eta)	Race (Eta)	Income (Tau$_b$)	Pol. Party I.D. (Eta)	Pol. Ideol. (Eta)
Direct interest issues														
Restrictions on smoking in public	.32***	.30***	.14***	.11***	−.23***	.07*	.17***	.04	.07**	.11***	.06†	.01	.04	.07
Sales taxes on tobacco products	.39***	.29***	.05	.19***	−.23***	.03	.05	.05†	.12***	.03	.16***	.12***	.12**	.06
Other policy/program issues														
Federal programs to aid tobacco farmers to switch	.16***	.09**	.05	.09**	−.17***	.03	.00	.03	.09**	.09**	.05	.03	.05	.06
Federal tobacco crop support programs	.11***	.08**	.14***	.10**	−.08**	.01	.16***	.02	.12***	.12***	.17***	.13***	.10*	.06
Complete ban on cigarette advertising[a]	.28***	.40***	.16***	.13**	−.17***	.14**	NA	.20***	.06†	.07	.02	.08*	.06	.06
Ethical/legal issues														
Right to smoke in public places	.28***	.29***	.11***	.11**	−.21***	.11***	.09**	.09**	.06*	.09**	.05	.00	.13***	.03
Morally wrong to profit from tobacco business	22***	.31***	.10**	.08*	−.11***	.12***	.04	.11***	.06*	.04	.04	.08**	.06	.07

NOTE: Analysis *N*s ranged from 773 to 946 cases when using data for both states.

[a] North Carolina data only; *N*s ranged from 400 to 463.

† $p \leq .10$. * $p \leq .05$. ** $p \leq .01$. *** $p \leq .001$.

MULTIVARIATE ASSOCIATIONS

All dependent variables are metric, but both metric and nonmetric levels are represented among the independent variables. Therefore, analysis of covariance, wherein the influence of each factor and covariate was assessed after controlling for the effects of all others in the equation, was used for the final analysis of each dependent variable.[7] The results displayed in table 3 point to several important conclusions.

First, variance explained (R^2) ranges from 6% to 26%, with at least one-fifth of the variance explained in four of the seven dependent variables. Not shown in the table is the fact that, for each of the seven dependent variables, the collective influence (main effects) of the factors and covariates in the equation is highly significant ($p \leq .001$).

Second (to elaborate upon a point partially developed in the discussion of table 1), the grand means generated by the covariance procedures also are informative. (The range of possible scores used for each of the first four dependent variables was -1 to $+1$, and was -2 to $+2$ for each of the final three dependent variables.) Specifically, the grand means reveal the absence of a consistent pro- or antismoking/tobacco position: there is overall support for smoking restrictions/prohibitions and for programs to help tobacco farmers switch to other crops, but also for the view that people have the right to smoke in public and to profit from tobacco products. Respondents are almost evenly divided concerning changes in cigarette sales taxes, federal crop support programs, and a complete ban on cigarette advertising.

Third, approximately half (34) of the 65 *apparent* significant influences on the seven dependent variables retain their significance after analysis by the covariance technique. Also, nearly half (16) of these 34 remaining significant influences are concentrated among the 3 current direct interest variables, a finding that supports Green and Gerken's results.[8]

Fourth, among the remaining self-interest variables, only two of the four other tobacco-related variables are significantly associated with as many as three of the dependent variables. The four significant asso-

---

[7] See SPSS Inc. 1986 (451–62) for a description of the analysis of covariance procedure, options, and statistics that enable simultaneous controls for all other factors and covariates in the equation, as well as the production of a multiple classification analysis table that includes R^2 and adjusted beta coefficients.

[8] The analysis shown in table 2 was also conducted without including "Smoke Bother," to examine if hostility toward smoking might color responses that were made by some respondents to the items of this scale. Deletion of "Smoke Bother" yielded only modestly different results, with one exception. Specifically, the influence of "Smoke Status" upon "complete ban on cigarette advertising" was dramatically elevated from nonsignificance (.08 in table 2) to high significance (.19***). Our interpretation is that it is not simply being a nonsmoker—but, rather, being a "bothered" nonsmoker—that explains a substantial proportion of the variance for this issue.

Table 3. Smoking Issue Opinions by Measures of Self-Interest and Sociodemographic Status: Multivariate Associations, 1984–85 (Adjusted Beta Coefficients)

			Measures of Self-Interest												
			Current Direct Interest Variables			Other Tobacco-related Variables				Measures of Sociodemographic Status					
Smoking Issue Opinions	R^2	Grand Mean	Smoke Status	Smoke Bother[a]	Profit from Smoke	Smoke Health Problem	Best Friends Smoke[a]	Parent Smoked	State	Age[a]	Educ.	Gender	Race	Income	Pol. Party I.D.	Pol. Ideol.
Direct interest issues																
Restrictions on smoking in public	.20	0.4	.19***	.21***	.12***	.01	−.10**	.03	.12***	—[b]	.05	.05	.06	—	—	—
Sales taxes on tobacco products	.23	0.2	.29***	.17***	.04	.08*	−.08*	.02	.05	.05	.04	—	.07*	.07	.06	—
Other policy/program issues																
Federal programs to aid tobacco farmers to switch	.06	0.6	.09*	.03	.06	.06	−.12**	.05	.02	—	.06	.06	—	—	—	—
Federal tobacco crop support programs	.13	0.1	.11**	.04	.13***	.05	−.04	.04	.15***	—	.07	.12***	.16***	.04	.06	—
Complete ban on cigarette advertising[c]	.26	−0.1	.08	.36***	.12**	.04	−.09	.08	NA	.19***	.05	—	—	.09*	—	—
Ethical/legal issues																
Right to smoke in public places	.21	−0.6	.14***	.24***	.10**	.02	−.14***	.10**	.10**	.06	.02	.00	—	—	.15***	—
Morally wrong to profit from tobacco business	.15	−0.7	.10**	.31***	.08*	.00	−.01	.08*	.05	.10**	.03	—	—	.07*	—	—

NOTE: Analysis of covariance with each factor and covariate adjusted for all others; analysis *N*s ranged from 705 to 868 cases when using data for both states.

[a] Covariate.

[b] Variable not entered into equation due to nonsignificant bivariate results.

[c] North Carolina data only, $N = 399$.

* $p \le .05$. ** $p \le .01$. *** $p \le .001$.

ciations involving the number of one's best friends who smoke are consistently inverse, and Illinois residents consistently demonstrate higher means (more anti-tobacco views) in all three significant associations involving "State."

Finally, most (17) of the 25 significant bivariate associations involving measures of sociodemographic status washed out, and none of these variables claim more than 2 of the 8 significant associations that remain. This general tendency is especially noteworthy for education. All seven of its significant bivariate associations washed out, a highly unusual finding for this typically powerful predictor of attitudinal and behavioral differences.

Discussion

We extended the Green and Gerken study in several ways. Whereas their data came from the west coast (California), ours came from the midwest (Illinois) and a southeastern tobacco state (North Carolina). Our instrument included several additional smoking policy items and a broader range of both individual and group-level self-interest measures. We tested the self-interest/policy preference linkage at the bivariate level, as did Green and Gerken, but also tested these associations at the multivariate level through an analysis of covariance.

Our findings at the bivariate level replicate—and at the multivariate level further support—those from the Green and Gerken study. In short, our data demonstrate that self-interest deriving from smoking status and "degree of bother" have significant effects on respondents' opinions about both public smoking restrictions and taxation. Going beyond the Green and Gerken analysis, we also found that these same two independent variables are somewhat related to our other policy/program issues and strongly related to both of our ethical/legal issues.

Both significant bivariate associations between "Smoke Bother" and the two federal programs to support farmers washed out when analysis was elaborated by the covariance procedure. We attribute this change to the possibility that the latter programs are not as concrete to individuals as either public smoking restrictions or tobacco sales taxes, which would be daily annoyances to smokers, or to a complete ban on cigarette advertising or either of the ethical/legal issues, which likely would have clear and direct salience to those respondents bothered by smoke.

The significant influence of self-interest is extended by our addition of a third current direct interest item: profit from smoke. It also relates significantly to most (5) of the 7 dependent variables analyzed in this report. In fact, the 3 current direct interest variables collectively account for 16 of the 34 significant influences that persist after conducting an analysis of covariance, and in every instance where one of these 3 variables exhibits a significant influence, it is in the direction that supports the earlier findings by Green and Gerken. It is always the nonsmokers, the higher scorers on the "Smoke Bother" scale, and the "No" respondents to the profit from smoke question, who demonstrate the more anti-tobacco (or less pro-tobacco) mean in any comparison.

Two other self-interest variables not considered by Green and Gerken also are moderately informative. The four significant associations found for "Best Friends Smoke" are consistently inverse, and Illinois residents consistently demonstrate higher means (more anti-tobacco views) in all three significant associations involving "State." Since the findings for state persist after controlling for other significant sociodemographic variables, we attribute the more pro-tobacco views of North Carolinians to the fact of their residence in "the tobacco state." Several sociodemographic variables not included by Green and Gerken also occasionally demonstrate significant influence on one or more dependent variables (i.e., one instance each for smoke health problems, gender, and political party; two instances each for parent smoked, age, race, and income).[9]

POLICY IMPLICATIONS

Green and Gerken assert that, unlike other social or economic issues such as jobs programs, smoking issues produce a strong linkage between self-interest and policy preference. Their explanation for this difference between smoking and other social issues is that the costs and benefits of public smoking restrictions and tobacco sales taxes are clear and salient. Our results support and extend their findings.

Among all of the independent variables that we added to the Green and Gerken study, another measure of self-interest—profiting from tobacco—is the strongest predictor of the smoking issue attitudes that we studied. Clearly, those respondents who personally profit from the manufacture or sale of tobacco products have the highest economic stakes in tobacco policy and are more likely to view most smoking issues from a pro-tobacco perspective. In contrast, those individuals who do not smoke—and/or are bothered by the smoke from those who do—have the strongest anti-tobacco stands on most of the smoking policy controversies that we examined. Self-interest, whether measured by "Smoke Status" or by "Smoke Bother," is less clearly a predictor of issues concerning government support of the tobacco farmer—possibly due, in part, to the abstract nature of these issues and to their lower salience among the majority of respondents. However, economic self-interest—as measured by "Profit from Smoke"—does conform to the self-interest hypothesis (i.e., it is significantly correlated with favoring increases in federal aid to farmers who grow and sell tobacco).

---

[9] Other recent research efforts report that smoking prevalence and behavior vary by group (Marin, Perez-Stable, and Marin 1989; Novotny et al. 1988; Pierce et al. 1989; Rogers and Crank 1988).

Appendix: Study Variables

(Except as noted, all percentages are based on a total *N* of 946—488 NC cases plus 458 IL cases.)

Dependent Variables

DIRECT INTEREST ISSUES

	Favor (+1)	*Oppose (−1)*	*DK*[a]
Would you favor or oppose a law requiring non-smoking areas—			
. . . in restaurants?	78%	19%	3%
. . . in government buildings like post offices?	65	29	6
. . . in offices or buildings where people work?	69	27	4
Would you favor or oppose a law prohibiting all smoking—			
. . . in grocery stores?	60	35	4
. . . in some hotel/motel rooms reserved for non-smokers?	73	22	4
. . . on airline flights lasting less than two hours?	66	30	4

[a] DK/NA responses were coded as missing values and those cases excluded from bivariate and multivariate analysis using pair-wise deletion.

Restrictions on Smoking in Public Scale (for analysis, recomputed to range from −1 to +1)

Sum of restrictions favored or opposed:	*Opposed*											*Favored*		
	−6	*−5*	*−4*	*−3*	*−2*	*−1*	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*DK*
	7%	6%	1%	0%	6%	1%	7%	2%	9%	1%	4%	16%	40%	0%

		Increased	*Left Same*	*Decreased*	*DK*
Should sales taxes on cigarettes be . . .	(+1 / 0 / −1)	32%	46%	18%	4%

OTHER POLICY/PROGRAM ISSUES

		Increased	*Left Same*	*Decreased*	*DK*
How about the federal money that supports farmers who grow and sell tobacco?	(−1 / 0 / +1)	22%	38%	30%	10%
What about programs to help tobacco farmers switch to growing other crops?	(+1 / 0 / −1)	69	18	7	5

		Strongly Agree	*Agree*	*Disagree*	*Strongly Disagree*	*DK*
[NC respondents, only] There should be a complete ban on cigarette advertising.	(+2 to −2)	10%	33%	42%	10%	5%

ETHICAL/LEGAL ISSUES

		Strongly Agree	*Agree*	*Disagree*	*Strongly Disagree*	*DK*
A person should have the right to smoke in public.	(−2 to +2)	11%	60%	23%	3%	3%
It is morally wrong for a person to profit from the growing, manufacturing, or selling of tobacco products.	(+2 to −2)	13	59	18	3	7

Independent Variables

CURRENT DIRECT INTEREST VARIABLES

Smoke Status

70%	Nonsmoker	("never smoked" cigarettes steadily for at least 3 months, plus "former smokers" who ever smoked for at least 3 months but not in the prior week)
29%	Smokers	("light," "medium," and "heavy" smokers combined)
1%	No answer	

How often has being around cigarette smoke caused you to have—	*Never*	*Sometimes*	*Often*	*Always*	*DK*
. . . a headache?	47%	35%	11%	5%	1%
. . . breathing problems?	53	29	9	8	1
. . . an eye irritation?	29	42	17	11	1

Smoke Bother Scale	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*DK*
How often bothered	20%	16%	19%	16%	8%	7%	6%	3%	1%	2%	2%

	Yes	*No*	*DK*
This last year, did you or a close relative receive any money from the growing, storing, processing, or selling of tobacco products?	9%	90%	1%

OTHER TOBACCO-RELATED VARIABLES

	Yes	*No*	*DK*
Have you or any relative ever had a serious health problem due to smoking?	27%	70%	3%

	0	*1*	*2*	*3*	*DK*
Think of your three best friends. How many of them smoke?	30%	24%	18%	27%	1%

	Yes	*No*	*DK*
Did either of your parents ever smoke while you were growing up?	69%	30%	1%

MEASURES OF SOCIODEMOGRAPHIC STATUS

State	North Carolina	52%	Illinois	48%			DK/NA	0%
Age	Mean	44						
Education	Hi. school or less	58%	Some col. or more	41%			DK/NA	1%
Gender	Male	48%	Female	50%			DK/NA	2%
Race	White	78%	Nonwhite	21%			DK/NA	1%
Total household income last year	< $20,000	37%	$20,000+	55%			DK/NA	8%
Party I.D.	Republican	28%	Independent	27%	Democratic	38%	DK/NA	7%
Pol. ideology	Liberal	17%	Moderate	37%	Conserv.	34%	DK/NA	12%

References

Green, Philip, and Ann Elizabeth Gerkin. 1989. "Self-Interest and Public Opinion toward Smoking Restrictions and Cigarette Taxes." *Public Opinion Quarterly* 53:1–16.

Marin, Gerardo, Eliseo J. Perez-Stable, and Barbara Vonoss Marin. 1989. "Cigarette Smoking among San Francisco Hispanics: The Role of Acculturation and Gender." *American Journal of Public Health* 79:196–99.

Novotny, Thomas E., Kenneth E. Warner, Juliette S. Kendrick, and Patrick L. Remington. 1988. "Smoking by Blacks and Whites: Socioeconomic and Demographic Differences." *American Journal of Public Health* 78:1187–89.

Pierce, John P., Michael C. Fiore, Thomas E. Novotny, Evridiki J. Hatziandreu, and Ronald M. Davis. 1989. "Trends in Cigarette Smoking in the United States: Educational Differences Are Increasing." *Journal of the American Medical Association* 261:56–60.

Rogers, Richard G., and John Crank. 1988. "Ethic Differences in Smoking Patterns: Findings from National Health Interview Surveys." *Public Health Reports* 103:387–93.

SPSS Inc. 1986. *SPSSX User's Guide*. 2d ed. New York: McGraw-Hill.

MEASURING HAPPINESS IN SURVEYS: A TEST OF THE SUBTRACTION HYPOTHESIS

ROGER TOURANGEAU,
KENNETH A. RASINSKI, and
NORMAN BRADBURN

Abstract Responses to an item on general happiness can change when that item follows one on marital happiness. Asking about marital happiness first sometimes reduces reported levels of overall happiness. This reduction may result from a change in the interpretation of the general happiness item. According to this hypothesis, when the general item follows the item on marital happiness, respondents take the item to ask about aspects of their lives other than their marriages—in effect, the respondents subtract their (mostly happy) marriages in answering the general item. The study reported here tests this "subtraction" hypothesis by asking versions of the general happiness item that correspond to the different interpretations. A version of the general item that asked about general happiness "aside from your marriage" yielded responses that were quite similar to those given to the standard item when it followed the item on marital happiness. Another version that asked about general happiness "including your marriage" elicited responses quite similar to those elicited by the standard item when it preceded the marital happiness item. However, reanalysis of the studies that originally demonstrated the impact of the order of the two happiness items casts doubt on the subtraction hypothesis and related models as explanations of the earlier findings.

The literature on survey context effects suggests that earlier questions can affect answers to later ones through a number of distinct mechanisms (Schuman and Presser 1981; Smith 1982; Strack and Martin

ROGER TOURANGEAU is Vice President at CODA, Inc. KENNETH RASINSKI is a Survey Director at the National Opinion Research Center (NORC) and NORMAN BRADBURN is NORC's Director. The authors thank Tom Smith and Howard Schuman, who carried out the reanalyses reported here. Their cooperation—and promptness—are greatly appreciated. The data collection described here was supported by a grant from the NORC Director's Fund.

Public Opinion Quarterly 55:255–266 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-02/$2.50

1987). In their recent review, Tourangeau and Rasinski (1988) argued that earlier questions can change how respondents interpret later items, what information they retrieve in formulating their answers, and how they make the judgments required by later items.

One survey context effect is widely believed to reflect processes operating as the item is interpreted. It involves a question asking about general happiness ("Taken altogether, how would you say things are these days? Would you say that you are very happy, pretty happy, or not too happy?"). This item has figured in some landmark studies of happiness and psychological well-being (e.g., Bradburn 1969). However, answers to the item can vary depending on whether it precedes or follows a similar item on marital happiness ("Taking things all together, how would you describe your marriage? Would you say that your marriage is very happy, pretty happy, or not too happy?"). Schuman and Presser (1981) reported a contrast effect in which respondents were significantly *less* likely to describe themselves as very happy when they had first answered the question on marital happiness. Complicating matters, Smith (1982) found the opposite—respondents were significantly *more* likely to say they were very happy when they had first answered the marriage question.

Of the two findings, Smith's is the less puzzling. Most respondents (typically, almost two out of three) report that they are very happily married. It is easy to see why reminding respondents of their mostly happy marriages just before asking them about their overall happiness will produce higher levels of reported happiness. The item on marital happiness should increase the chance that respondents will recall their marriages and take them into account in evaluating their overall happiness (McClendon and O'Brien 1988; Smith 1985; Strack, Martin, and Schwarz 1988). Such retrieval-based assimilation effects have been commonly observed, both with items on happiness and with other attitude items (e.g., Tourangeau et al. 1989a, 1989b). Though readily explained, the assimilation effect is likely to be weak. Even without the reminder provided by the preceding item, marriage is likely to be a salient consideration for most respondents as they assess their overall happiness; because the prior item on marital happiness will only affect those respondents who would not have taken their marriages into account anyway, the impact is likely to be small. Consistent with this analysis, efforts to replicate Smith's findings, such as the one reported by Turner (1984), have generally found nonsignificant effects in the direction of assimilation.

Schuman and Presser's finding demonstrates that, though assimilation may be the rule, there are definite exceptions. Their results raise two questions: Why does the contrast effect occur? What determines which effect occurs? In answer to the first question, several investigators have suggested that the contrast effect reflects differences in how the general happiness item is interpreted under the two question-orders. Bradburn (1982), for example, argued that when the general item comes first, respondents may see it as referring to all aspects of their lives including their marriages, but that when it comes second, they see it as referring to aspects of their lives other than their marriages. Because they exclude or "subtract" their marriages in evaluating their overall happiness, respondents who answer the general happiness item after the item on marriage may give a different answer than they would have given had the general item come first. Others have proposed similar accounts of this or related context effects (e.g., Kalton, Collins, and Brook 1978; Schuman and Presser 1981).

Researchers have sometimes linked the subtraction hypothesis to more fundamental processes. Bradburn argued that respondents may consider it redundant to base their answers to both happiness items on evaluations of their marriages. Tourangeau (1984; see also Strack, Martin, and Schwarz 1988; Tourangeau and Rasinski 1988) proposed that this hypothesized tendency to avoid redundancy may reflect an even more basic principle that governs behavior in conversations—namely, the principle that each contribution should be informative (Grice 1975; Haviland and Clark 1974). If this principle is at work in survey interviews, respondents may interpret the general item to be asking for new information—that is, about areas of their lives not already covered by the preceding item on marital happiness.

The subtraction account suggests some variables that may affect whether assimilation or contrast will occur. For example, summaries are not always redundant, particularly when many or diverse specifics have been touched on. As Tourangeau and Rasinski point out, a long list of prior specifics "may encourage respondents to interpret the general item as a summary of the particulars rather than as a residual category" (1988, 303). Thus, when the marriage and general happiness items are part of a longer series of related items, contrast effects may be less likely. The number of prior specific items may provide a clue about why Smith found an assimilation effect but Schuman and Presser found contrast (Schwarz, Strack, and Mai 1991). In the 1980 General Social Survey (GSS) on which Smith's results are based, the two happiness items followed five questions about satisfaction in other life domains; Schuman and Presser's study did not include the satisfaction items.

Despite its popularity, the evidence in favor of the subtraction account for the contrast effect has mostly been quite indirect. Smith (1982) noted that the subtraction hypothesis implied that the effect of the preceding marital happiness item should depend on how respondents answered that question—that is, the effect of excluding marriage from the judgment of overall happiness should depend on whether the marriages being left out are, on the average, happy or unhappy ones. Strack, Martin, and Schwarz (1988) went further, arguing that the difference in interpretation should affect the correlation between responses to the two happiness items; the correlation should be reduced when marriage is excluded from the evaluation of general happiness.

But the central claim of the subtraction hypothesis—that the difference in question order can produce a difference in the interpretation of the general happiness item—has only recently been tested. The subtraction hypothesis states that, when the general happiness item follows the item on marital happiness, respondents interpret it to mean something like "Aside from your marriage, how happy are you?" but that when it comes first, respondents interpret it to mean something like "Including your marriage and other aspects of your life, how happy are you?" If this is true, then it should be possible to duplicate the results obtained under the two question orders merely by rephrasing the general happiness item to make the intended interpretation clear and explicit. This is what a recent study by Schwarz and his colleages attempted to do (Schwarz, Strack, and Mai 1991).

It is possible that this study does not fully capture the essential features of the original studies. First of all, Schwarz and his co-workers made several changes in the wording and format of the happiness items. Their items asked about satisfaction instead of happiness and provided an 11-point response scale instead of the usual 3-point scale. The biggest wording change involved replacing the item on marriage with an item that covered dating relationships as well; this allowed unmarried respondents to take part in the study. However, in a similar study that examined order effects on items asking about satisfaction in several life domains, McClendon and O'Brien (1988) reported differences between married and unmarried respondents—only the married respondents showed assimilation effects. Schwarz and his colleagues do not report the results separately for married and unmarried respondents or say what proportion of their sample was married, so we cannot assess whether including unmarried respondents affected their results.

Second, the respondents in the new study were a convenience sample of volunteers who filled out written questionnaires. The change from personal interviews in the earlier studies to self-administered questionnaires in the new one may have altered the results; elsewhere, Schwarz and Strack have argued that self-administered questionnaires reduce or eliminate question order effects (Bishop et al. 1988; Schwarz et al., in press). Third, although the new study shows clear evidence of assimilation effects, the evidence for contrast is much weaker. Respondents who received the version of the general item that explicitly instructed them to leave out their marital or dating relationship did not show a significantly reduced correlation between the two happiness items relative to the control group, although the difference was in that direction.

We carried out a study that avoided some of these problems; our study reproduces the main features of the original studies more exactly than the study by Schwarz, Strack, and Mai (1991). It uses telephone interviews carried out by experienced interviewers with a randomly selected sample of married respondents. Our study also includes conditions that retain the exact wordings of the original items. Finally, we carried out a detailed comparison of our results with those of the original studies and a recent replication on the 1987 GSS. This comparison proved to be quite revealing.

Method

DESIGN

The study was a split-ballot experiment with four groups. Respondents in two groups got the standard wording of the general happiness item ("Taken altogether, how would you say things are these days? Would you say that you are very happy, pretty happy, or not too happy?"). In one of these standard wording groups, the general happiness item preceded the item on marital happiness; in the other, it followed the item on marital happiness. The two standard wording groups thus replicate the conditions reported by Schuman and Presser (1981) and Smith (1982). The final two groups received variant wordings for the general happiness item. For one group, the item read: "Taking things all together, including your marriage and other important aspects of your life, how would you say things are these days? Would you say that you are very happy, pretty happy, or not too happy?" For the other group, the general item read: "Aside from your marriage, how would you say things are these days? Would you say that you are very happy, pretty happy, or not too happy?" In both variant wording groups, the general happiness item followed the item on marital happiness.

All four groups got the standard wording of the marital happiness item ("Taking things all together, how would you describe your marriage? Would you say that your marriage is very happy, pretty happy, or not too happy?"). In all four, the two happiness items were the last questions in the interview and immediately followed an item that asked about the respondent's marital status; twelve items on welfare and abortion preceded the items on marital status and happiness.

SAMPLE

The experiment reported here was conducted as part of a larger study in which a sample of households in Chicago was contacted by telephone and individuals were selected for interview; those who completed an initial interview were recontacted and interviewed over the telephone a second time. The happiness items were administered during the second interview.

We drew the sample for the initial interview by randomly selecting telephone numbers from the Chicago directory and replacing the final digit of each number with a random digit. Interviewers called this initial pool of telephone numbers and contacted a total of 1,481 households; calls to an additional 570 numbers were never answered. The selection of a respondent within the household was nonrandom, with interviewers asking for any available adult male or, when no male was available, any adult female. The initial interview, which concerned attitudes on abortion and welfare, was completed by 599 respondents.

Respondents to the initial interview were recontacted by telephone three weeks later. Four hundred ninety-nine of the original respondents also completed the reinterview; 228 of them were married. Of the 228, 48.7% were men and 51.3% were women; 55.3% said they were white, 27.4% Black, 5.8% Hispanic, and 11.4% classified themselves in other race or ethnic categories. Their ages ranged from 18 to 87 years, with a mean of 47 years.

PROCEDURE

Ten experienced telephone interviewers carried out both the initial interviews and the reinterviews. Each interviewer received a batch of reinterview questionnaires, with the different versions in a random order. By working their way through the questionnaires, the interviewers randomly assigned reinterview respondents to one of the four experimental groups.

Results

The hypotheses can be summarized quite simply. The correlation between responses to the two happiness items should be relatively high when the general item preceded the item on marital happiness or when the wording of the general item made it clear that respondents were to take their marriages into consideration. The correlation should be relatively low when the general item followed the item on marital happiness or when the wording of the general item made it clear that respondents were to exclude their marriages from consideration.

Table 1 presents the correlations for the four groups. As is apparent, they conform quite closely to the hypothesized pattern. The two correlations that are supposed to be high are considerably higher than the two that are supposed to be low (.545 and .524 vs. .277 and .267). A contrast that compares the first two of these correlations with the last two reveals a significant difference ($z = 2.32$; $p < .05$). Table 1 also shows the significance of the differences between pairs of these correlations; both pairwise comparisons are marginally significant.

Although the correlations between the items vary across the four groups, the marginals on the general happiness item do not. The proportion of respondents describing themselves as "very happy" was 34.0% among those who got the standard item after the marital happiness item, 40.0% among those who got the standard item before the marital happiness item, 33.3% among those who got the "aside from your marriage" revised wording, and 33.9% among those who got the "including your marriage" revised wording. None of the differences in the marginals for the general happiness item are significant.

Discussion

In terms of the correlations between the two happiness items, respondents who got the revised item asking them about their lives aside from their marriages show results very similar to those who received the standard general happiness item after the item on marital happiness. Those who got the revised item asking about their lives including their marriages were quite similar to those who got the standard item first. These results are certainly consistent with the view that the revised items reflect the different interpretations that respondents give to the standard item under the different question orders. What is particularly striking about the findings is how closely the correlations for the two revised groups parallel those for the two standard wording groups.

Our findings differ somewhat from the results in the comparable conditions of the study by Schwarz, Strack, and Mai (1991). First, in our study, when the standard general item followed the item on marital happiness, there was a low correlation between the two ($r = .28$); in their study, the corresponding correlation was quite high ($r = .67$). Schwarz and his colleagues included another condition that yielded results more like ours. In that condition, the general item followed the specific item and the two items were explicitly linked by a joint introduction; within this group, the correlation ($r = .18$) was comparable to the value of .28 that we found. As it happens, our study included

Table 1. Responses to the General Happiness Item and Correlation with Marital Happiness

Group	N	"Very Happy" Percentage	z	p	Correlation with Marital Happiness	z	p
Standard wording							
Marital happiness first	53	34.0	0.66	n.s.	.277	−1.69	<.10
General happiness first	60	40.0			.545		
Revised wording							
"Aside from your marriage"	54	33.3	−0.07	n.s.	.267	−1.59	<.12
"Including your marriage"	59	33.9			.524		

NOTE: All significance tests are two-tailed.

such a lead-in to the happiness items; just before these items, there was a short introduction that read "Now for a couple of items on an unrelated topic." Even such a vague introduction apparently tended to heighten the relationship between the two items, affecting interpretations of the general item.

A second difference between the results of the two studies involved the correlation between the two happiness items when the general item came first. In our study, this correlation was relatively high (r = .55) when the standard version of the general happiness item was used; in the study by Schwarz and his colleagues, the corresponding correlation was relatively low (r = .32). This difference may reflect the reduced importance or accessibility of dating relationships as compared to marriages. The specific item in the study by Schwarz and his colleagues covered both marriage and dating; ours asked only about marriages. The difference in the correlations in the two studies may also reflect the fact that our study included an item to determine whether the respondent was married; this question came right before the two happiness items. The question on marital status may have temporarily heightened the accessibility of the respondent's marriage, increasing its impact on answers to the general happiness item. Data from the earlier studies suggest that both explanations have some merit. For the conditions in which the general item came first, Smith and Schuman and Presser report correlations between the two happiness items that are somewhat lower than ours but somewhat higher than the one reported by Schwarz and his co-workers (r = .39 in Smith 1982, and r = .49 in Schuman and Presser 1981).

Despite these superficial differences, the results of our study are basically consistent with those of Schwarz and his colleagues. The key question that remains is whether either study, both of which found differences in correlations, sheds any light on the original studies, both of which found differences in the overall direction of responses to the general item. If the contrast effect reported by Schuman and Presser results from the reinterpretation of the general happiness item when it follows the marital happiness item, there should be a correlational effect as well as a directional one; as in table 1, the correlation between the two happiness items should be reduced when the general item comes second. Similarly, if the assimilation effect reported by Smith results from the increased accessibility of marital happiness when the item about marriage comes first, this should heighten the correlation between the two items. Were there such correlational effects in the earlier studies?

Apparently not. Table 2 presents the results of a reanalysis of the data from the earlier studies and from a recent replication conducted as part of the 1987 GSS. None of the studies showed a significant difference between the correlations obtained under the two question orders. The results shown in table 2 do not depend on the measure of association used; the pattern remains the same whether correlation coefficients, simple or multiple regression coefficients, or gammas are used to measure the association between the two happiness items.

There *is* one situation in which both the subtraction and accessibility hypotheses predict directional rather than correlational effects. That situation is when everyone answers the marital happiness item in more or less the same way. The absence of variability in responses to the specific item will guarantee a low correlation with the general item, regardless of question order. Further, because all respondents will be subtracting (or including) material with similar implications for the general happiness question, the specific item should affect answers to the general item in the same way for all respondents, producing a directional effect. The data in table 2, however, are not consistent with this hypothesis either. The correlations there are not particularly low, ranging from the high .30s to the low .50s. These values are closer to the high correlations in table 1 (and in the study by Schwarz and his colleagues) than to the low ones. To put it another way, the correlation between responses to the two happiness items remains relatively high when the respondents are supposed to be discounting their marriages (as in Schuman and Presser 1981) and it gets no higher when the respondents are supposed to have been reminded of their marriages by the preceding item on marital happiness (as in Smith 1982).

Table 2. Responses to the General Happiness Item and Correlation with Marital Happiness

Study and Group	N	"Very Happy" Percentage	z	p	Correlation with Marital Happiness	z	p
Schuman and Presser 1981							
Marital happiness first	172	37.2	3.14	<.01	.533	0.56	n.s.
General happiness first	201	53.2			.490		
Smith 1982							
Marital happiness first	613	42.1	−2.22	<.05	.417	0.44	n.s.
General happiness first	268	34.3			.390		
General Social Survey 1987							
Marital happiness first (total)	519	40.5	−1.37[a]	n.s.	.373	0.56[a]	n.s.
Satisfaction items before happiness items	261	39.5	−0.46[b]	n.s.	.418	1.25[b]	n.s.
Satisfaction items after happiness items	258	41.5			.323		
General happiness first	265	35.5			.409		

NOTE: The five satisfaction items preceded the happiness items in both conditions of Smith 1982; they were omitted from Schuman and Presser 1981.

[a] Significance tests measure the marital happiness first condition vs. the general happiness first condition.

[b] Significance tests compare the two satisfaction conditions.

Thus, our study and the one by Schwarz and his colleagues do not appear to capture whatever processes produced directional effects in the original studies. Like Schwarz and his colleagues, we found correlational, but not directional, effects; the earlier studies found directional, but not correlational, effects. Although the results of the newer studies support the idea that respondents may reinterpret apparently redundant items (though perhaps only when the items are explicitly linked by a joint introduction), such reinterpretation does not seem to have produced the contrast effect observed by Schuman and Presser. Similarly, although there is considerable evidence that earlier items can increase the accessibility of material that is then more likely to be used to answer later questions, this does not appear to explain the assimilation effect found by Smith. Moreover, the results of the 1987 GSS experiment indicate that the placement of the satisfaction items had little impact on responses to the two happiness items. This variable does not seem to determine whether assimilation or contrast effects are observed (cf. Schwarz, Strack, and Mai 1991; Tourangeau and Rasinski 1988).

In short, our conclusions must remain negative. The subtraction hypothesis does not appear to explain the contrast effect involving the happiness items. Similarly, the more complicated model offered by Schwarz and his colleagues, which incorporates the subtraction hypothesis as a special case, also offers no explanation of the original findings; further, that model's explanation of the assimilation effects found in the 1980 GSS appears to be contradicted by the data shown in table 2. Unfortunately, although we may now be clearer about what is *not* causing the original effects, we are still unsure about what is causing them.

References

Bishop, G., H. Hippler, N. Schwarz, and F. Strack 1988. "A Comparison of Response Effects in Self-Administered and Telephone Surveys." In *Telephone Survey Methodology*, ed. R. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, and J. Waksberg, 321–40. New York: Wiley.

Bradburn, N. 1969. *The Structure of Psychological Well-Being*. Chicago: Aldine.

Bradburn, N. 1982. "Question-Wording Effects in Surveys." In *Question Framing and Response Consistency*, ed. R. Hogarth, 65–76. San Francisco: Jossey-Bass.

Grice, H. 1975. "Logic and Conversation." In *Syntax and Semantics*, vol. 3: *Speech Acts*, ed. P. Cole and T. Morgan, 41–58. New York: Seminar Press.

Haviland, S., and H. Clark. 1974. "What's New? Acquiring New Information as a Process in Comprehension." *Journal of Verbal Learning and Verbal Behavior* 13:512–21.

Kalton, G., M. Collins, and L. Brook. 1978. "Experiments in Wording Opinion Questions." *Journal of the Royal Statistical Society*, ser. C, 27:149–61.

Jabine, T., M. Straf, J. Tanur, and R. Tourangeau. 1984. *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*. Washington, DC: National Academy Press.

McClendon, M., and D. O'Brien. 1988. "Question-Order Effects on Subjective Well-Being." *Public Opinion Quarterly* 52:351–64.

Schuman, H., and S. Presser. 1981. *Questions and Answers in Attitude Surveys: Experiments in Question Form, Wording, and Context*. New York: Academic Press.

Schwarz, N., F. Strack, H. Hippler, and G. Bishop. In press. "The Impact of Administration Mode on Response Effects in Survey Measurement." *Applied Cognitive Psychology*.

Schwarz, N., Strack, F., and H. Mai. 1991. "Assimilation and Contrast Effects in Part-Whole Question Sequences: A Conversational Logic Analysis." *Public Opinion Quarterly* 55:3–23.

Smith, T. 1982. "Conditional Order Effects." GSS Technical Report no. 33. Chicago: NORC.

Smith, T. 1985. "Unhappiness on the 1985 GSS: Confounding Change and Context." GSS Technical Report no. 34. Chicago: NORC.

Strack, F., and L. Martin. 1987. "Thinking, Judging, and Communicating: A Process Account of Context Effects in Attitude Surveys." In *Social Information Processing and Survey Methodology*, ed. H. Hippler, N. Schwarz, and S. Sudman, 123–48. New York: Springer-Verlag.

Strack, F., L. Martin, and N. Schwarz. 1988. "Priming and Communication: The Social Determinants of Information Use in Judgments of Life Satisfaction." *European Journal of Social Psychology* 18:429–42.

Tourangeau, R. 1984. "Cognitive Science and Survey Methods." In *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*, ed. T. Jabine, M. Straf, J. Tanur, and R. Tourangeau, 73–100. Washington, DC: National Academy Press.

Tourangeau, R., and K. Rasinski. 1988. "Cognitive Processes Underlying Context Effects in Attitude Measurement." *Psychological Bulletin* 103:299–314.

Tourangeau, R., K. Rasinski, N. Bradburn, and R. D'Andrade. 1989a. "Belief Accessibility and Context Effects in Attitude Measurement." *Journal of Experimental Social Psychology* 25:401–21.

Tourangeau, R., K. Rasinski, N. Bradburn, and R. D'Andrade. 1989b. "Carryover Effects in Attitude Surveys." *Public Opinion Quarterly* 53:495–524.

Turner, C. 1984. "Why Do Surveys Disagree? Some Preliminary Hypotheses and Some Disagreeable Examples." In *Surveying the Subjective Phenomena*, vol. 2, ed. C. Turner and E. Martin, 159–214. New York: Russell Sage.

THE INFLUENTIALS: BACK TO THE CONCEPT OF OPINION LEADERS?

GABRIEL WEIMANN

Abstract The concept of opinion leadership has been related to a lingering theoretical and methodological debate. The present study examines a new measure to identify opinion leaders (the Strength of Personality Scale) developed by the Allensbach Survey Center in Germany. The results of applying this measure in Germany and Israel allowed the testing of its validity and efficiency. The findings, however, do not suggest that the influentials identified by this scale are opinion leaders according to the original conceptualization, but fit better the more sophisticated characterization of opinion leadership that stemmed from modifications of the original concept of opinion leaders.

> The mass do not now take their opinions from dignitaries in Church or State, from ostensible leaders, or from books. Their thinking is done for them by men much like themselves, addressing or speaking in their name, on the spur of the moment. . . .
>
> —John Stuart Mill, *On Liberty*[1]

About 100 years after John Stuart Mill had first suggested the importance of personal influence and the role of influentials, several studies provided the first empirical evidence of these concepts (Berelson, Lazarsfeld, and McPhee 1954; Katz 1957; Katz and Lazarsfeld 1955; Lazarsfeld, Berelson, and Gaudet 1948; Merton 1949). The role of

---

GABRIEL WEIMANN is Associate Professor of Communication and Chairman of the Department of Sociology at the University of Haifa, Israel. The author gratefully acknowledges the important contributions of Elisabeth Noelle-Neumann, Hans Eysenck, Hans-Bernd Brosius, and Robert Eisinger. The research was supported by the Alexander von Humboldt-Stiftung, Bonn, Germany.

[1] The full sentence goes on to indicate that this flow of interpersonal influence is mediated by the newspapers.

Public Opinion Quarterly 55:267–279 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-02/$2.50

interpersonal relations in the flow of information and influence, as revealed by these studies, caused a growing interest in personal networks and in key positions in these networks. Consequently, many studies focused on influential positions or the "opinion leaders." Several hundred opinion leadership studies (see reviews by Rogers and Shoemaker [1971] and Köpller [1984]) have attempted to identify the characteristics of opinion leaders in terms of demographic and socioeconomic variables, media exposure, social positions, and personality traits.[2]

The measurement of opinion leadership was subjected to empirical criticism and debate (e.g., Gitlin 1978; Katz 1980, 1989; Lin 1973; Robinson 1976; Weimann 1982). This criticism, combined with the need for more sophisticated and validated measurement of opinion leadership, has led to the development of new measurement procedures. One of the recent attempts is the research project conducted by Noelle-Neumann and the German survey research center, the Allensbach Institut für Demoskopie, on a new instrument to identify opinion leaders, the "Strength of Personality Scale" (PS scale).

The present study is based on applying this scale to different samples in two societies and testing its validity by various measures, including a comparison with measures of actual flow of information and influence based on sociometric mapping of the personal network of a small community.

Method

The first scale to be called the "Strength of Personality Scale" (*Personalichkeitsstarke*) emerged from testing numerous questionnaire items related to self-perceived levels of personal influence. (The German news magazine, *Der Spiegel,* initiated the development of the new measure. Interested in what they referred to as "the active consumers who set standards in their community," managers of *Der Spiegel* challenged the Allensbach Institut with the task of developing an instrument that would identify the influentials.) These early scales were tested and refined after years of pretests with a variety of samples (see

---

[2] Several studies focused on the level of education as it related to opinion leadership (e.g., Kingdon 1970; Reynolds and Darden 1971; Saunders, Davis, and Monsees 1974; Summers, 1971), or on gender (e.g., Booth and Babchuck 1972; Eurich 1977; Kingdon 1970) or social class (e.g., Booth and Babchuck 1972; Summers 1971). Other studies tried to relate personality traits such as conformity, social insecurity, responsibility, motivation, anxiety, and emotional stability to the status of opinion leaders (e.g., Brett and Kernalegün 1975; Robertson and Myers 1969; Saunders, Davis, and Monsees 1974; Schrank and Gilmore 1973; Summers 1971).

Noelle-Neumann 1983, 413–15; 1984 and 1985). The final scale was established after a factorial reduction of a 34-item questionnaire administered to a representative sample of 3,542 residents of Germany. The resulting scale includes 10 items that were later weighted according to their part-whole correlations with the total scale, with the scores on this scale ranging from 75 to 149 (see table 1). The respondents were divided into four approximate quartiles, thus yielding four levels of personal strength (for a detailed description of the questionnaire, including wording of all items and the respondents' breakdowns, see Noelle-Neumann 1983).

The procedure used by the Allensbach Institut to validate the sample involved the comparison of the scale ratings with the interviewer's impression (whether the respondent "radiates strength and energy" as opposed to "being boring," whether the respondent gave the impression of "being a model for others" and whether he or she appeared "self-assured" during the interview). The comparison of the results revealed that interviewers' ratings and the PS scale ratings are highly correlated (Noelle-Neumann 1985, 11). However, this test is weakened by the possibility that the interviewers were influenced by the respondent's self-assessment. The need to validate the scale by an "external" criterion led to the present study.

The scale was administered to two Israeli samples. The first (*N* = 650) was randomly drawn from the adult, Jewish population of Israel. The respondents were interviewed at home with a prestructured questionnaire that included 34 statements, among which were the final 10 PS scale items (the respondent had to choose those cards that carried statements suited to his or her self-assessment, thus dichotomizing the answers). Additional questions included sociodemographic characteristics, measures of community participation, media exposure, functioning as advice givers in various areas, patterns of sociability, and social involvement. This first sample was used to examine the applicability of the PS scale to another society and to study cross-cultural differences.

The second sample was an Israeli kibbutz community (*N* = 270) that underwent a sociometric mapping of its personal communication network and of the flow of information and influence in this network (for a detailed review of these studies, see Weimann 1982, 1983; also see Granovetter 1981, 29–30). A sociometric mapping of the personal communication network was obtained by interviewing every member of the community: each respondent was asked to list his or her conversational ties with other members (unlimited number of choices) and was instructed to rate the strength of each tie by three scales (importance, frequency of contacts, and tenure of the tie), thus operationalizing Granovetter's definition of ties' strength. Every tie was rated by

Table 1. The Personality Strength (PS) Scale and Weighting

Item	Weight Yes	Weight No
1. I usually count on being successful in everything I do	13	7
2. I am rarely unsure about how I should behave	14	7
3. I like to assume responsibility	15	7
4. I like to take the lead when a group does things together	17	8
5. I enjoy convincing others of my opinions	15	7
6. I often notice that I serve as a model for others	16	8
7. I am good at getting what I want	14	7
8. I am often a step ahead of others	18	9
9. I own many things others envy me for	15	9
10. I often give others advice and suggestions	12	6
Maximum score	149	
Minimum score		75

NOTE: The weighting procedure is detailed in Noelle-Neumann 1985 (5–7).

the two connected persons and the scores given by each member of the dyad to the same tie were summed, thus adding the element of mutuality. The sociometric data was arranged in a who-to-whom matrix with each entry ($N = 2{,}511$ ties) representing a communication tie, characterized by its strength. The entire matrix was subjected to a clique identification procedure, obtained by a computerized sequential reordering of the matrix (see Richards 1975, 1977; a detailed example is provided in Rogers and Kincaid 1981, 163–82).

To study the actual flow of use of a tie for communicating information, six different items were used, two from each category: "general news," "consumer information," and "gossip" (for a description of the items, see Weimann 1982). None of these items was publicized formally, but were disseminated a month prior to the study by means of interpersonal communication. To trace the flow, each member was asked whether he or she was aware of each of the items and to report on the source of his or her knowledge (who told him or her). Two decision-making items were used to measure the flow of influence. These were two issues later put to vote in the *assefa* (the general assembly): a decision to allow young members, after their military service, to spend a year outside the kibbutz, and a decision to accept or reject the application of a candidate for full membership. Each member was asked about consulting, seeking advice, or being influenced when making up his or her mind regarding these issues. The entries about flow of information and influence were added to the sociomatrices, in order to identify network positions and their activity in the flow of information and influence. Finally, we applied the PS scale and cross-classified this measure with the communication network attributes (position and activity) for each subject.

Findings

The first step involved the cross-cultural comparison of the scale's results by applying several statistical procedures to the data collected in the German sample (N = 3,542) and the first Israeli sample (N = 650). A factor analysis of the responses to the 34-item questionnaire was performed for both samples (using a principal components solution with varimax rotation). It revealed that, in both samples, the same 10 items used for the PS scale emerged as strongly loaded on one factor, accounting for about 41% of the variance in each sample. This suggests stability across social and cultural settings. A split-half reliability test for both samples yielded coefficients of .78 and .76 for the German and Israeli samples, respectively. When the 10-item scale itself is factor analyzed, two distinct factors emerge: a principle one accounting for 75% of the variance (72% in the Israeli sample) that includes seven items (1, 2, 3, 4, 5, 7, and 10) and a second factor that includes three items (6, 8, and 9). The first factor seems to measure internal sources of influenceability, whereas the second describes "external" origins, derived from comparison with other people (Noelle-Neumann 1985).

The respondents in both samples were divided into four groups according to the quartile scores on the PS scale.[3] Following the procedure of the Allensbach studies, the relationships between the PS and sociodemographic characteristics were examined. The breakdowns of the German and the first Israeli samples, obtained by cross-classifying PS groupings with various sociodemographic factors, indicate that PS is related to various sociodemographic variables, and these relationships are very similar for both populations. In terms of gender, there is a sharp difference between males and females: males score higher on PS (male-to-female ratio in the "strong" PS level is 1.47 in the German sample and 1.88 in the Israeli one). PS is also unevenly distributed by age groups: the highest PS scores are more frequent in the 30–39 age group (in both samples) while the "weak" level is more frequent in the 60 plus age group (23% of this age group are rated "weak" in the German sample and 48% in the Israeli one).

---

[3] The distributions of the two samples were different. Consequently, the quartiles do not have the same ranges.

The strongest distinction is revealed by SES classification: a person in the higher SES levels has more than three times as much chance to be ranked as "strong" PS as a person in the lower SES levels. The weaker PS persons are much more frequent among the lower SES. Again, this pattern is almost identical for both populations. This relationship does not indicate direction of causality: opinion leaders may achieve their leadership through social acknowledgment of their education, income, occupation, and popularity, but they may also achieve higher social positions because of their stronger personality.[4]

Although the PS scale yields similar results in two different societies, this does not indicate its validity. Is it really a measure of influenceability? The application of the PS scale to a social network that was mapped by its communication links, personal positions, and flow of information and influence may provide the "external" criteria for testing the scale's validity.

The communication matrix included 2,511 entries connecting the 270 members of the kibbutz. The analysis of the 270 × 270 matrix structure according to Richard's procedure (see Rogers and Kincaid 1981, 163–82) revealed 16 cliques. Using the clique identification, we computed the network position of each individual in his or her clique, by the number of links with members of his or her clique. The division into four quartiles (of number of ties) resulted in four levels of network positions, from "marginals" to "centrals." This measure, originally gathered for a network analysis (Weimann 1982, 1983), was compared with the results of applying the PS scale to the same respondents. The two measures were highly correlated: a correlation coefficient of .54 was found between the number of communication links (in the entire network) and the PS measure, and this correlation was somewhat higher (.59) when relating PS score to the number of communication links within the clique (the difference between the coefficients is not statistically significant). Individuals with high strength of personality are better linked to other individuals in the community and especially to others in their social group. This relationship is clearer when the groupings according to the two scales are cross-classified, as shown in table 2.

Individuals with central positions in their groups are clearly those with higher levels of PS: 58% of those who are centrally positioned in the communication network rated as "strong" PS, while only 3% of the "marginals" rated as "strong" PS individuals. The same applies

---

[4] In the Israeli study, we included a question about military service (obligatory for all Israeli men and women). The analysis revealed that the military ranking is highly correlated with the PS grouping ($\gamma = .75$). Thus, 54% of the officers were rated in the "strong" PS, while 61% of the nonofficers were rated in the "weak" PS.

Table 2. Network Positions and Strength of Personality (in percentages)

	Strength of Personality				
Network Position	Strong	Above Average	Moderate	Weak	*N*
Central	58	17	14	11	65
High medium	31	28	25	16	68
Low medium	15	28	31	26	68
Marginal	3	13	23	61	69

NOTE: $\chi^2 = 85.23$, $df = 9$, $p < .0001$; $\gamma = 0.686$, $p < .0001$.

to other network positions, with decreasing PS score according to lower network position (61% of the "marginals" in their groups are rated as "weak" on the PS scale, while only 11% of the "centrals" belong to this category of PS). The strong gamma coefficient of .68 indicates that these two indicators are interrelated. The strong relationship with sociometric centrality may indicate that the PS scale is not exclusively a personality measure but an instrument that also reflects social position. The early studies of Allensbach found the same pattern (though not by sociometric measures). Thus, for example, the German surveys revealed that high PS respondents report having more friends and acquaintances (see Noelle-Neumann 1985, esp. table 10: 67% of the "strong" PS group reported having many friends and acquaintances, compared to only 49% of the "above average," 37% of the "moderate," and 20% of the "weak"). While one may question the validity of such self-assessments, the data from the sociomatrix obtained by a mapping of actual communication links in the Israeli kibbutz community reveal the same relationship.

The final and crucial validity test of the PS scale is its ability to predict communicative and influential behavior. To do so, measures of the actual flow of various information and decision-making items in the kibbutz community were compared with the PS scores. As the mapping of the communication activity was based on dyadic measurement ("who-to-whom"), the comparison of communicative activity of individuals according to their PS level was possible. To do so, we cross-classified each tie activated for the flow of a specific item with the communicator's PS level. Table 3 presents the communicative activity of the four PS groups in terms of their relative share in the flow. The activity rates are the percentages of the ties activated by each level of PS in the flow of six information items and two decision-

Table 3. Communicative Activity by Strength of Personality

Item	Ties Activated[a]							Average PS			
	Strong 1	2	3	Weak 4	N	χ^2	p	Communicator	Non-communicator	t^b	p
Gossip											
Item 1	22	31	32	15	425	7.7	n.s.	116.8	114.5	1.08	n.s.
Item 2	24	36	28	22	382	5.6	n.s.	117.1	115.7	0.65	n.s.
General news											
Item 1	34	26	24	16	501	6.5	n.s.	120.2	109.8	4.19	<.001
Item 2	38	22	22	18	474	9.4	<.05	116.2	109.1	2.18	<.01
Consumer information											
Item 1	41	28	19	12	375	18.8	<.001	133.4	102.8	12.24	<.001
Item 2	39	31	17	13	381	17.6	<.01	130.9	104.6	10.52	<.001
Decision making											
Item 1	48	31	19	2	346	45.2	<.001	137.9	101.5	14.57	<.001
Item 2	46	29	18	7	366	33.2	<.001	138.3	100.6	15.09	<.001

[a] Percentage of ties activated by a communicator whose strength of personality is 1, 2, 3, or 4.

[b] t-test conducted for the differences between the averages under the null hypothesis of no significant difference.

making items. The higher the percentage, the more active the individuals from the specific PS group were. In other words, a bigger share of the communication flow is related to communicators from a specific PS level.

The proportions of ties activated by various levels of PS vary across items. However, except for the flow of gossip, the most active communicators are those with higher levels of PS. The majority of ties activated for the flow of news, consumer information, and influence originated from individuals with stronger PS. The rates of activation decreased for lower levels of PS, indicating the strong relation between communicative activity and strength of personality (except for gossip). This relation is evidently more significant for decision-making items than for information items. In the case of influence flow (the decision-making items), almost half of the flow originated from individuals with "strong" PS (while only 2%–7% of this flow originated from the "weak" PS group). In other words, the influentials, identified by actual activity, are mainly those rated as "strong" on the PS scale.

Another comparison presented in table 3 is of PS averages: for each of the items studied, we compared the average PS for the "communicators" (those who were listed by the respondents as the source of information or influence) and the "receivers" (those who were not listed as sources). The pattern revealed by the averages confirms our previous finding: the PS averages are significantly higher for the communicators for most of the information items and for both decision-making items. This difference is largest for the flow of influence and smallest (and actually not significant) for the flow of gossip. Thus, individuals with stronger PS are more active as communicators of news and consumer information and are clearly the influentials for decision-making items.

Conclusions

From a theoretical perspective, the identification of the "influentials" should be related to the concept of opinion leaders. The criticism of the concept of opinion leaders has focused mainly on its methodological deficiencies and its emphasis on the two-step flow model. However, a better measurement procedure and a more realistic presentation of the influentials, regardless of the media effects controversy, should cause a revision of this concept. Successfully identifying the influentials, validating the measurement by "external" criteria, and pointing to the role of these influentials in the flow of interpersonal communication and influence suggests a "return to the opinion leadership idea." However, as the findings of both the German and the Israeli studies suggest, this revised measure of personal influence is independent of the flow of mass communication.[5]

Opinion leadership was originally conceptualized as a combination of personal and social factors. As Katz (1957) noted, being an influential is related to three attributes: (1) the personification of certain values (or "who one is"); (2) competence ("what one knows"); and (3) strategic social location ("whom one knows"). The present study combines these characteristics by relating them to actual communicative activity. The individuals identified as "strong" on the PS scale combined personal traits, competence, and social position in their personal network, enabling them to influence others. Thus, it is not merely a unidimensional measure, but a combination of personal traits with social network positioning ("It is, however, not enough to be a person whom others want to emulate, or to be competent. One must also be accessible" [Katz 1957, 74]). The PS scale's strong correlation with network positions and communicative activity is not only a methodological validation of an empirical instrument, but also provides a theoretical linkage between the personal and social dimensions of the opinion leaders concept.

In spite of the similarities between the influentials as identified by the PS scale and the opinion leaders, these terms are not fully interchangeable. Four differences emerge from a first comparison. (1) While opinion leaders were found by Katz and Lazarsfeld (1955) to be fairly evenly distributed among the different social strata, the influentials (as measured by PS scaling) were concentrated in higher SES levels. However, later studies of opinion leadership also revealed this unequal distribution (e.g., Booth and Babchuck 1972; Hamilton 1971; Kingdon 1970). (2) The opinion leaders were found to differ from one subject to another with very little overlap, while the influentials were found to be active in several areas with considerable overlap (see Noelle-Neumann 1985, 47). Again, later studies of opinion leaders revealed some overlap, even in the data used by Katz and Lazarsfeld (Marcus and Bauer 1964). (3) Opinion leaders were found to be considerably more exposed to mass media, while this tendency was not significantly related to the influentials. In fact, it appears that the influentials rely on personal sources and their special pattern of media consumption, not merely on higher media exposure.[6] This notion of personal influence on opinion leaders was suggested by Katz when revising his two-step flow model: "Opinion leaders, though more exposed to the media than nonleaders, nevertheless reported personal influence as the major factor in their decisions" (Katz 1957, 76). As to the pattern of media consumption, later studies of opinion leadership related different quality and form of media consumption to opinion leaders (e.g., Hamilton 1971; Levy 1978; Ostlund 1973). (4) While the opinion leadership concept was based on the leader-follower dichotomy, PS scaling reveals that influenceability is, rather, a continuous variable. Again, the possibility of a continuum was suggested by later studies of opinion leadership (e.g., Lin 1973). Thus, it appears that the influentials are not the opinion leaders according to the original definition but fit better the more sophisticated characterization that stemmed from the growing criticism of early studies on opinion leadership.

---

[5] The Allensbach Studies as well as the Israeli study found no significant differences between PS levels in terms of media exposure (e.g., exposure to television and radio was almost identical for all the PS groups; see Noelle-Neuman 1985, 27).

[6] For example, there was no significant difference between PS levels in terms of watching television or listening to the radio (Noelle-Neumann 1985, table 18) when asked about leisure activities. In the Israeli sample, the higher levels of PS did not report higher exposure to newspapers, television, or radio. However, some indication of the

From a methodological perspective, the findings of the present study highlight the efficiency of the new measure. The relationships between the PS measure and real activity as communicators and influentials provides the needed validation. The universality of the scale, when applied to two different societies, further contributes to its usefulness. Further studies should focus on the relationships between social status and PS (revealed to be stronger than suggested by early studies of opinion leaders) and measure the weight of personal traits and social characteristics that constitute the status of an influential.

Finally, the identification of influentials by means of the PS scale has important implications for practical-minded scholars of politics, marketing, advertising, and public opinion in general. The identification of influentials and the study of their sources of information and decision-making processes may have substantive value to those who combine social research with persuasive communication. The development of the PS scale is a crucial first step for such an integration.

References

Berelson, Bernard, Paul F. Lazarsfeld, and W. N. McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign.* Chicago: University of Chicago Press.

Booth, Alan, and Nicholas Babchuck. 1972. "Informal Medical Opinion Leadership among the Middle Aged and Elderly." *Public Opinion Quarterly* 36:87–94.

Brett, Joyce E., and Anne Kernaleguen. 1975. "Perceptual and Personality Variables Related to Opinion Leadership in Fashion." *Perceptual and Motor Skills* 40:775–79.

---

different quality of media consumption was revealed by the Allensbach studies, especially when PS ratings were related to reading certain newspapers and specific contents of the papers; see Noelle-Neumann 1983, tables 27l, 27m, and 27n.

Eurich, Carl. 1977. "Beinflussungsstrategien zur Weiterentwicklung des Meinungsfuhrerkonzeptes." *Zeitscrift für Markt, Meinungs and Zukunftsforschung* 20:42–85.

Gitlin, Todd. 1978. "Media Sociology: The Dominant Paradigm." *Theory and Society* 6:206–53.

Granovetter, Mark S. 1981. "The Strength of Weak Ties: A Network Theory Revisited." Revised version of a paper presented to the International Communications Association Meetings (May, 1980), Acapulco.

Hamilton, Herbert. 1971. "Dimensions of Self-Designated Opinion Leadership and Their Correlates." *Public Opinion Quarterly* 35:266–74.

Katz, Elihu. 1957. "The Two-Step Flow of Communication: An Up-to-Date Report on an Hypothesis." *Public Opinion Quarterly* 21:67–78.

Katz, Elihu. 1980. "On Conceptualizing Media Effects." *Studies in Communication* 1:119–41.

Katz, Elihu. 1988. "Communication Research Since Lazarsfeld." *Public Opinion Quarterly* 51:S25–S45.

Katz, Elihu, and Paul F. Lazarsfeld. 1955. *Personal Influence*. Glencoe, IL: Free Press.

Kingdon, John W. 1970. "Opinion Leaders in the Electorate." *Public Opinion Quarterly* 34:256–61.

Köpller, Karlfritz. 1984. *Opinion Leaders*. Hamburg: Heinrich Bauer Verlag.

Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet. 1948. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. New York: Duell, Sloan and Pearce.

Levy, Mark R. 1978. "Opinion Leadership and Television News Uses." *Public Opinion Quarterly* 42:402–6.

Lin, Nan. 1973. *The Study of Human Communication*. Indianapolis: Bobbs-Merrill.

Marcus, Alan S., and Raymond A. Bauer. 1964. "Yes: There Are Generalized Opinion Leaders." *Public Opinion Quarterly* 28:628–32.

Merton, Robert K. 1949. "Patterns of Influence." In *Communication Research*, ed. P. Lazarsfeld and F. Stanton, 180–219. New York: Harper and Brothers.

Noelle-Neumann, Elisabeth. 1983. *Spiegel Dokumentation: Personlichkeitsstarke*. Hamburg: Spiegel Verlag.

Noelle-Neumann, Elisabeth. 1984. *The Spiral of Silence*. Chicago: University of Chicago Press.

Noelle-Neumann, Elisabeth. 1985. "Identifying Opinion Leaders." Paper presented at the 38th ESOMAR Conference, Wiesbaden.

Ostlund, Lyman. 1973. "Interpersonal Communication Following McGovern's Eagleton Decision." *Public Opinion Quarterly* 37:601–10.

Reynolds, Fred D., and William R. Darden. 1971. "Mutually Adaptive Effects of Interpersonal Communication." *Journal of Marketing Research* 6:441–50.

Richards, William D. 1975. *A Manual for Network Analysis*. Institute for Communication Research, Stanford University. Mimeo.

Richards, William D. 1977, "Network Analysis Methods: Conceptual and Operational Approaches." Paper presented at the fourth annual colloquium on Social Networks, University of Hawaii, Honolulu.

Robertson, Thomas S., and James H. Myers. 1969. "Personality Correlates of Opinion Leadership and Innovative Buying Behavior." *Journal of Marketing Research* 6:155–67.

Robinson, John. 1976. "Interpersonal Influence in Election Campaigns: Two-Step Flow Hypothesis." *Public Opinion Quarterly* 40:304–19.

Rogers, Everett M., and Laurence B. Kincaid. 1981. *Communication Network: Toward a New Paradigm for Research*. New York: Free Press.

Rogers, Everett M., and Floyd F. Shoemaker. 1971. *Communication of Innovations: A Cross-Cultural Approach*. New York: Free Press.

Saunders, John, Michael J. Davis, and David M. Monsees. 1974. "Opinion Leadership in Family Planning." *Journal of Health and Social Behavior* 15:217–27.

Schrank, Holly L., and Lois D. Gilmore. 1973. "Correlates of Fashion Leadership: Implications for Fashion Process Theory." *Sociological Quarterly* 14:534–43.

Summers, John O. 1971. "Generalized Change Agents and Innovativeness." *Journal of Marketing Research* 8:311–15.

Weimann, Gabriel. 1982. "On the Importance of Marginality: One More Step into the Two-Step Flow of Communication." *American Sociological Review* 47:764–73.

Weimann, Gabriel. 1983. "On the Strength of Weak Conversational Ties." *Social Networks* 5:245–67.

IN MEMORIAM

I. A. Lewis, 1921–1990

He was born Irwin Albert Lewis, but no one ever called him anything but "Bud." When he died of cancer last August in Los Angeles, he was 68. Capping a distinguished career in survey research, he had been polling director for the *Los Angeles Times* since 1979.

Bud had a fertile imagination and an expansive personality. He was often argumentative, but always lovable. One friend, in describing Bud's command of centerstage at any gathering, said, "He can sure fill a room." He was a man who had his own style in everything he did, whether it was polling or his study of architecture or just enjoying the "best" restaurant in town.

Bud practiced his polling art for leading news organizations in the United States. He became director of polling for NBC News in 1966 and later was director of its election unit. He left NBC and joined CBS News as director of surveys in 1976, during the early days of the CBS News/*New York Times* Poll. He worked for the Roper Organization the following year, before starting the *Los Angeles Times* poll.

He was equally proud of his election coverage and polling abroad, most recently in Nicaragua. Earlier, he did work in Puerto Rico and the Philippines, from which he narrowly escaped after Ferdinand Marcos discovered that Bud was conducting election polls that would be a fair reflection of the true will of the people.

He attended Princeton University and was graduated with honors in English. During World War II, he lived in a palace in India while serving in Air Force Intelligence. After the war, his career took him to Germany as a bureau chief for the Armed Forces Network, where, among other things, he produced a musical show for American troops. He also worked for United Press International and the Voice of America.

Starting in 1956, he was a writer for Dave Garroway on the original "Today" show at NBC. He was replete with stories about Garroway and J. Fred Muggs, the chimpanzee Garroway used on the show. It was Bud's job to write dialogue for Garroway, who conversed with the chimp.

WARREN J. MITOFSKY

Public Opinion Quarterly 55:280 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-02/$2.50

THE POLLS—A REVIEW

WHICH SIDE ARE YOU ON? THE 1990 NICARAGUAN POLL DEBACLE

PETER V. MILLER

The celebrated "polling debacle" in the 1990 Nicaraguan election brought a variety of responses: for some pollsters, there was bewilderment, consternation, reassessment, apology, and defense. For others, the reactions included schadenfreude, self-righteousness, and excoriation. Some viewed the widespread mismatch between poll predictions and election results as a failure of ethnocentric North American polling methods or the product of a naive misconception of what polls can achieve in formerly totalitarian societies. Others argued that the flawed predictions were more the result of a highly volatile electorate than of ill-conceived or botched polling procedures.

Such points were made in an atmosphere that—like the Nicaraguan election itself—was highly politicized and contained no small measure of intrigue. Observations of the polls were made in an epistemological "twilight zone" where one's political position was all-important and polls were propaganda weapons. Neutral observation was virtually impossible; observers were branded with a political label by default—"If you're not on my side, you must be the enemy." Partly for this reason, there was much "missing data" on how preelection polls in Nicaragua were actually carried out. Attempts to understand the differences among polls were hampered by the refusal of some polling organizations to share information with "the enemy" and by the inherent difficulties in monitoring far-flung field operations.

PETER V. MILLER is Associate Professor of Communication Studies and Journalism and Director of the Institute for Modern Communication at Northwestern University. This review would not have been possible without the sponsorship and documentation provided by the Commission on Nicaraguan Preelection Polls. In particular, I want to thank William Bollinger, who coordinated the commission's work and was primarily responsible for collecting the information on the polls. I am indebted to Nancy Belden for additional information on Itztani and DOXA polls, to Richard Soudriette of the International Foundation for Electoral Systems for information on VIA CIVICA, and to Janice Ballou for information on Borge y Asociates. The review has benefited from the careful scrutiny of Nancy Belden, Katherine Bischoping, Carlos Denton, William Millard, and the Poll Review editors. Finally, I wish to thank Rebecca Jane Sliva-Brockington for valuable research assistance and translations of Spanish documents.

Public Opinion Quarterly 55:281–302 © 1991 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/91/0055-02/$2.50

One cannot hope to uncover "truth" in such circumstances. In lieu of that lofty aim, I propose, in this review, to describe and evaluate various answers to the question of what happened in the Nicaraguan preelection polls. I will examine major hypotheses offered to explain the differences among polls and between the polls and the election results. I will discuss the work of the Commission on Nicaraguan Preelection Polls, the entity through which most of what is known about the polls has come to light.[1] Finally, I will offer some views about the significance of the Nicaraguan poll "failure" and subsequent attempts to make sense of it.[2]

Synopsis of the Polling Problem

Violeta Chamorro of UNO, an amalgam of fourteen parties united in opposition to the Sandinista regime in Nicaragua, was elected president with nearly 55% of the vote on February 25, 1990. Daniel Ortega of the FSLN (Sandinista) party was second in the presidential balloting with nearly 41% of the vote. Eight other parties divided the remainder of the presidential ballots.

Compared to the election results, only half of the 29 polls taken within a year of the election were "directionally correct"—showing UNO with a higher vote intention percentage (see table 1). Perhaps more striking, the polls maintained the same level of disparity—with half favoring UNO and half favoring the FSLN—throughout 1989 and right up to the final weeks of the campaign. Of the five published polls fielded in the month of February, 1990, three favored UNO and two supported the Sandinistas. The lack of convergence of the polls on some similar picture of reality as election day approached is a primary example of the "polling problem" in the Nicaraguan election.

Further, the disparity between the "UNO-leading" and "FSLN-leading" polls was not small. If the "undecided" respondents in each poll are allocated equally to the two major contenders, it is possible to compare the difference between the estimates from the two groups of polls and the final vote percentages. The 14 published polls taken within a year of the election showing the Sandinistas in the lead underestimated the eventual UNO vote by an average of about 20 percentage

---

[1] I am a member of the commission, and visited Nicaragua under its aegis. The "raw material" for the review was obtained by members of the commission or its staff, consisting of methods documentation, data analyses, pollsters' press releases and reports to clients, as well as observations and interviews by the commission. I am responsible for any errors in the synthesis presented here.

[2] This review went to press in early 1991, approximately one year after the Nicaraguan election.

Table 1. Published Nicaraguan Preelection Polls, March, 1989, through February, 1990

Date	Sponsor[a]	Direction[a]	Field Work[a]	Population Sampled	N	Lead	Percentage Deviation from Final Vote[b]
3/28–4/6/89	Morales Foundation	Unknown	Unknown	Selected Municipalities	2,800	UNO	−6
5/13–5/14/89	UCA	Itztani	Itztani	3 cities	1,500	FSLN	−29
6/27–7/6/89	*La Prensa* *El Nuevo Herald* *La Nacion*	Borge Y Asoc.	Borge Y Asoc.	Selected Munic.	1,220	UNO	+6
7/29–7/30/89	INOP	Itztani	Itztani	28 Munic.	1,500	FSLN	−14
8/8–8/11/89	Morales Foundation	Unknown	Unknown	4 cities	1,150	FSLN	−24
9/27–10/2/89	CARIN	Itztani	Itztani	17 Munic.	1,157	FSLN	−11
10/2–10/7/89	UCA	ECO	ECO	31 Munic.	4,114	FSLN	−16
10/12–10/21/89	Various U.S. & Central Am. clients	CID-Gallup	CID-Gallup	28 Munic.	956	UNO	−3
10/22–10/27/89	Univision	Bendixen-Schroth Consultaria 21	Logos	23 Munic.	1,129	FSLN (virtual tie)	−8

Table 1. *(Continued)*

Date	Sponsor[a]	Direction[a]	Field Work[a]	Population Sampled	*N*	Lead	Percentage Deviation from Final Vote[b]
10/23–10/30/89	(IPCE NED) *La Prensa* *El Nuevo Herald* *La Nacion*	Borge Y Asoc.	Borge Y Asoc.	15 Munic.	1,284	UNO	+6
11/13–11/19/89	UCA	ECO	ECO	National	3,681	FSLN	−18
11/7–11/15/89	Unknown	DOXA	DOXA	14 Munic.	800	UNO	−5
11/25–12/3/89	Hemisphere Initiatives	Greenberg-Lake	Itztani	National	971	FSLN	−16
11/28–12/5/89	Unknown	DOXA	DOXA	Unknown	1,200	UNO	−3
12/89	UCA	ECO	ECO	National	4,431	FSLN	−25
12/9–12/16/89	NORAD CINZONTLE	Unknown	Unknown	Women only	1,000	FSLN	−21
12/15–12/20/89	NED	VIA CIVICA	VIA CIVICA	36 Munic.	2,825	UNO	+7
12/18–1/15/90	IPCE (NED) *La Prensa* UNO Campaign	Borge Y Asoc.	Borge Y Asoc.	National	1,200	UNO	−3
1/90	Hemisphere Intitiatives	Greenberg-Lake	Itztani	National	901	FSLN	−23

1/15–1/19/90	UCA	ECO	ECO	National	4,545	FSLN	−20
1/10–1/17/90	Unknown	DOXA	DOXA	Unknown	1,080	UNO	−6
1/25–1/31/90	Unknown	DOXA	DOXA	Unknown	1,200	UNO	−15
1/21–1/25/90	NED	VIA CIVICA	VIA CIVICA	36 Munic.	2,825	UNO	+3
1/23–1/28/90	Univision	Bendixen-Schroth Consultaria 21	Logos	National	1,000	FSLN	−16
1/24–2/7/90	Unknown	Mora y Araujo, Noguera y Asociados	Mora y Araujo, Noguera y Asociados	National	1,118	UNO	−1
2/90	IPCE (NED) *La Prensa* UNO Campaign	Borge Y Asoc.	Borge Y Asoc.	Unknown	8,000		−1
2/90	UCA	ECO	ECO	National	2,357	FSLN	−22
2/11–2/15/90	Washington Post & ABC News	Belden & Russonello	Itztani	National	925	FSLN	−15
2/11–2/13/90	IPCE (NED) *La Prensa* UNO Campaign	Borge Y Asoc.	Borge Y Asoc.	Managua only	1,260	UNO	−5

SOURCE: Data compiled by the Commission on Nicaraguan Preelection Polls.

[a] Identity of sponsors, directing research firms, and fieldwork agencies is discussed in the text. NED stands for National Endowment for Democracy, a U.S. agency that gave funds to Nicaraguan opposition groups, e.g., VIA CIVICA.

[b] Deviation is calculated after allocating the "undecided" respondents equally to UNO and FSLN.

points, even in the final weeks of the campaign. The 14 published UNO-leading polls taken in the same period, in contrast, were—with one exception—within 7 percentage points of the final UNO vote percentage, even months before the election.

Finally, UNO- or FSLN-leading polls were often produced by firms, groups, or sponsors who were affiliated (or who were charged with being affiliated) with one of the major parties. For example, a number of the UNO-leading polls were produced by the polling firm of Victor Borge, a Costa Rican political advisor to UNO. On the other side, some of the largest FSLN-leading estimates were produced by a Nicaraguan firm, ECO, whose sponsor—the Central American University (UCA)—was regarded as Sandinista leaning, at least by members of the opposition. Some of the polls showing UNO with the biggest leads were conducted by VIA CIVICA, a group opposed to the Sandinistas and funded by the U.S. National Endowment for Democracy. Finally, some of the polls showing large FSLN leads—including a poll taken shortly before the election for the *Washington Post*—were conducted by the Nicaraguan firm Itztani, which had previously worked with Central American University and whose director had been a government employee under the Sandinistas (but had later been jailed by them).[3]

The main picture we have, then, is of two polling "camps" that consistently and widely diverged in their estimates of electoral preference throughout the year preceding the campaign and right up to the election.[4] When matched against the eventual election outcome, the estimates from one of the camps turned out to be right, accounting for the perplexed reaction on one side and self-righteousness on the other. As noted, if "undecided" respondents are allocated equally to UNO and FSLN, 13 of the 14 published UNO-leading polls conducted in the year preceding the election produced UNO vote intention estimates within 7 percentage points of the eventual election percentage. The average absolute deviation from the actual UNO vote share for all 14 polls is 5%, about the same as the average deviation noted by Crespi (1988) in his examination of 430 U.S. preelection polls. Thus, the UNO-leading polls, as a group, appear to have been reasonably "accurate," though generally not accurate enough to meet Crespi's empirically derived criterion of "high accuracy."[5] Equally interesting from the point of view of Crespi's analysis of U.S. polls, and in sharp contrast to his findings, is the fact that the relative closeness of the UNO-leading polls to the election results does not depend much on the proximity of the poll to election day.

---

[3] Marvin Ortega, Itztani's director, had been a part of the Sandinista revolution (as had Violeta Chamorro) and had worked in the government agriculture department. But he became a critic of the Sandinistas and was jailed by them for a time. Still, he was viewed by UNO partisans as allied with the government.

[4] There are a few exceptions to this general rule (table 1). Two polls were carried out by the Morales Foundation, one showing a UNO lead and one an FSLN lead. A Univision-sponsored pair of polls, designed by the Washington firm Bendixen-Schroth in collaboration with a Venezuelan firm, Consultants 21, and fielded by the Nicaraguan firm LOGOS, showed, first, a virtual tie between UNO and the FSLN and, later, an FSLN lead.

Similarly, the divergence of the FSLN-leading polls from the election results appears to hold up regardless of when they were taken. The estimates from the 14 published FSLN-leading polls shown in table 1 differ from the eventual election results by no less than 11 percentage points and by as much as 29%. The smallest deviation occurs in a poll taken some five months prior to the election.[6] Looking at the results by region of the country (in the few cases where this is possible), the FSLN-leading polls were closest to the election results in war-torn Region 6, where the Greenberg-Lake November and January polls showed Chamorro with a slight lead (she carried the region with 70% of the vote).

Searching for Answers

FRAUD

Some simple explanations for the wide divergence among the polls are readily apparent. Perhaps the simplest is that the FSLN-leading polls were never conducted at all, but "poll results" were manufactured during the election campaign by Sandinista supporters to advance the election effort. This explanation is favored by Victor Borge, the UNO political advisor who supervised a number of UNO-leading polls.[7] In an environment that was not conducive to easy independent verification of polling activity and a political context so bitterly polarized, it is easy to imagine that "poll results" might be invented for propaganda purposes.

---

[5] See Crespi 1988 (21–24) for details of his empirically derived accuracy classification of U.S. polls. We commonly speak of polls being accurate or inaccurate, depending upon the closeness of their vote intention numbers to election results. This is not to say, however, that a close match between a poll's vote intention percentages and the final vote necessarily implies better polling procedures, or vice versa.

[6] While the vote intention marginals of the two polling camps varied widely, it is reasonable to ask if the polls showed similar patterns of correlations between vote intention and various predictors. For the most part, this possibility cannot be examined because such relationship data for most of the polls, particularly the UNO-leading polls, is not available. The available data concerning polls from both sides (CID-Gallup, October; ECO, November; Greenberg-Lake, November and January) show some similarity in the correlations between gender, age, education, and vote intention. These show women somewhat more supportive of UNO or more undecided. They generally show a positive relationship between Ortega support and education. Finally, they show increasing undecided and decreasing Ortega support with age.

[7] Borge argued (in an interview with Janice Ballou [1990] of Rutgers University, a member of the Commission on Nicaraguan Preelection Polls) that at least some of the FSLN-leading polls were never conducted.

The difficulty with this explanation is that disclosure of methodological information is, as a general rule, much better for the FSLN-leading polls than for the UNO-leading polls. Firms that produced higher FSLN vote intention estimates were substantially more receptive to the inquiries of the Commission on Nicaraguan Preelection Polls, and provided much more detail about their work than did Borge, for example, who treated methodological details as trade secrets. In early February, I personally examined notarized sampling records and completed questionnaires at the polling firm ECO, which produced some of the poll numbers most favorable to the Sandinistas. Other members of the commission observed interviews being taken by employees of ECO and LOGOS, whose second poll for Univision showed a substantial Sandinista lead. In contrast, most of those who conducted UNO-leading polls refused even to meet with outside observers during the preelection period, and, during one commission visit to Managua, claimed that their polling data had been destroyed by the Sandinistas.[8]

This is not to say that falsification of data could not have occurred and may have occurred to some degree on both sides. Rather, those who produced polls showing a Sandinista lead were generally more willing to document that they actually did the work than were their UNO-lead counterparts.

SAMPLING

The second explanation for the disparity among the polls is that those who produced FSLN-leading estimates used polling methods that were systematically flawed. Under this explanation one might hypothesize that, though the work was carried out in good faith, the sampling and/or data collection methods employed by these pollsters led to seriously and consistently erroneous findings.

Random sampling error is not a credible explanation for the differences among the UNO-leading and FSLN-leading polls because the differences among the polls, and their sample sizes, are too large.[9] Sampling does provide a source of concern, however, because of the practical difficulties of executing an area probability sample in Nicaragua, a country where telephone penetration is low and face-to-face interviewing is required. Prior to the provision of voter registration information by the Supreme Electoral Council in late 1989, pollsters had to draw samples based on population estimates by INEC, the Nicaraguan central statistical agency. INEC, in turn, was hampered in its work by the fact that no national census of Nicaragua had been performed in nearly 20 years, an earthquake had destroyed Managua in the early 1970s (the city is still largely in ruins), and the country had been wracked by war for more than a decade. Thus, pollsters had to start with questionable information on the distribution of the population across the country. In addition, INEC maps were not always reliable. ECO, the polling firm employed by the University of Central America, prided itself on a staff of professional cartographers whose job was to make up for the deficiencies in government-provided maps.

---

[8] As Howard Schuman and I concluded our visit to Managua during the first weekend in February, *La Prensa* (1990) published a story claiming that its polling data and computing equipment had been sabotaged by the Sandinistas. The story went on to claim that the Commission on Nicaraguan Preelection Polls was a Sandinista-front organization. The commission's difficulties gaining information from polling groups that showed a UNO lead will be discussed subsequently.

After voter registration in October, 1989, the Supreme Electoral Council (CSE) provided information on the structure of the electorate, organized by voting districts (JRVs, or *juntas*). There were some 4,400 JRVs with an average of 350 voters each. Fifty-six percent of the JRVs were classified as "urban." The CSE published summary figures on the electorate, broken down by sex, age, "urban" and "rural," and region of the country.[10]

Maps of the JRVs were available from the council. The JRV sampling frame offered the promise of more standardization in sampling, as pollsters would be selecting among the same set of well-defined geographic areas. Still, there remained considerable leeway for polling organizations to draw samples shaped by subjectivity and bias, and the practical difficulties of executing a sample design in the field were very great. Maps of the JRVs, again, were not always reliable. Each organization had to address cartography problems, decide how to sample areas *within* the voting districts, and then dwellings and voters within those areas.

---

[9] In all, some 58,000 interviews were taken in the polls considered in table 1, if news reports and disclosures by pollsters are to be believed. Some individual polls had enormous *N*s—particularly those conducted by ECO for the Central American University—in order to measure vote intentions in National Assembly races. Even the smallest polls had nearly 1,000 respondents. The fact that each interview could be done for less than $5.00 meant that even the largest studies cost less than many national polls in the United States with far smaller sample sizes.

[10] The CSE analysis of the electorate, cross-tabulating registrations by sex and age, suggested to me that it would be possible to select a name sample from their records, obviating the need for area sampling, quotas, etc. On my visit to Managua, I asked if it were possible to interview a voter sample, and was told by Raul Obregon at ECO that it would be "too expensive" to do so. (See n. 9 for a discussion of interview costs. A voter name sample, with call backs and no substitutions, undoubtedly would have been more expensive than the quota method employed by ECO and others.)

Although each of these sampling stages could be carried out according to probability methods, the possibility of biased selection exists, particularly in the selection of dwellings and respondents. None of the polls appears to have employed strict probability sampling. Prospective voters were selected by various quota methods, trying to match the CSE's figures on sex and age; information on nonresponse for initially selected respondents is scanty. More generally, the possibility for interviewers to select those who appeared to have the "right attitude" while filling their quotas, or for respondents to refuse to talk with interviewers who appeared to be on "the other side" is quite real.

Further complicating the sampling picture, less than half of the 29 polls considered here purported to be "national." Many involved samples of municipalities, regional capitals, and so forth. Most of the "national" polls were conducted after voter registration information became available.[11] Ironically, 8 of the 11 or so polls that purported to sample the national population showed the FSLN leading, while the (eventually) more accurate UNO-leading polls were often based on samples of municipalities, or even of one municipality (Managua). One would reckon that sample designs that were based on probability selection of JRVs from all over the country were superior in coverage to samples of regional capitals. A partial explanation for the higher accuracy of UNO-leading municipality polls may lie in the national sample execution in the field, where, as noted, conditions for completing interviews were often very difficult and there was room for supervisor and interviewer discretion in filling respondent quotas. Further, though respondent refusals were reported anecdotally by several polling organizations to be small (less than 10%), records of nonresponse were generally not kept.

Thus, it is possible that sample execution played a role in the inaccuracy of the FSLN-leading polls. It is difficult to believe, however, that all, or even a large portion, of the difference between the UNO- and FSLN-leading polls can be attributed to this factor. There is no evidence, for example, that sample execution was any better in the UNO-leading polls, which were also based on quota sampling. In fact, the only available external evidence about how sampling was done in the field comes from limited observation of polls that produced FSLN-leading estimates. In these cases, observers saw no evidence of biased selection or substantial nonresponse.[12] A further limitation on the possible impact of biased respondent selection on vote intention estimates is that it would be difficult to perform unwitting—that is, not fraudulent—biased selection frequently enough to effect the magnitude of difference between the UNO- and FSLN-leading polls. These considerations suggest that sample execution may have played some role in the different vote intention estimates, but that other factors were necessary and probably more important.

---

[11] Hemisphere Initiatives claimed to have sponsored the first "truly national" poll, fielded by Itztani and using voter registration information for the sampling frame, in late November and early December, 1989. ECO, however, purported to have done a "national" poll earlier in November, using information from INEC as the sampling frame.

DATA COLLECTION

Another set of hypotheses concerns various aspects of measurement: for example, voter screens, the form of the vote intention question, and its placement in the questionnaires employed in different polls. Polls differed in how they identified "likely" voters (and some were taken long before the October, 1989, voter registration). Some polls asked respondents for their vote intention directly, while others used various forms of a self-administered "secret ballot." As is frequently the case in U.S. polls, some of the Nicaraguan efforts placed the vote intention question first in the questionnaire to avoid possible context effects, and some placed it after varying numbers of other questions. These measurement differences, if systematic, could play a major role in explaining the disparity between FSLN- and UNO-leading polls.

Before examining these matters directly, it is important to consider the meaning and context of the term, *vote intention* in the 1990 Nicaraguan election. First, democracy in Nicaragua has a tortured history, shaped by civil war, dictatorship, and U.S. intervention since the mid-nineteenth century. Elections have been staged during this period, but the idea of popular sovereignty has not carried as much weight as has the application of force. The 1990 election was the first in which all of the various political factions in the post-Somoza era were represented on the ballot, and it featured unprecedented monitoring by international bodies. Further, the "secret ballot" concept was so tenuous that the Supreme Electoral Council—the body charged with supervising the election—sponsored an advertising campaign through which the slogan, "Your ballot is secret" saturated the country. Elaborate and well-publicized measures to combat election fraud (including the presence of hundreds of international observers), while intended to reassure the voting populous, also signaled the unparalleled nature of this election.

---

[12] Again, the limited field observations are not proof that biased respondent selection did not occur, or that there were not substantial numbers of refusals; I merely report that the only available empirical evidence argues against these hypotheses.

In this unique environment, asking people to report for whom they would vote takes on a different character than in a "normal" poll. There were no "normal" polls in Nicaragua, as polling itself had only been allowed since 1988. Respondents literally may not have known how to respond to vote intention inquiries, since responding and voting were both so novel. Moreover, the prospect of actually being able to carry out the vote was not at all guaranteed.

Further complicating the vote intention report was the highly polarized political context in which polls were conducted. As I have noted, poll results were trumpeted in the unabashedly partisan press as part of election propaganda. Respondents may not have taken vote intention questions seriously, given the partisan lens through which polls were interpreted. They may have taken the "Your ballot is secret" slogan to heart, and treated poll responses as one thing and voting as another. We know that some people did this explicitly, refusing to answer the vote intention question on these grounds, but others may have given a vote intention response that they did not mean. Instead of drawing a connection between their poll response and their vote, they may have sought to report what they thought the interviewer would like to record.

Finally, given the history of conflict and intimidation in the country, respondents may have been afraid to say how they would vote. Both sides reported many acts of intimidation to UN monitors, ranging from armed violence to "voter registration checking" by party workers. Some "polls" were carried out by partisan activists on both sides. For any or all of these reasons, measuring vote intention during the Nicaraguan election campaign was certain to be extraordinarily difficult.

It is also possible that voters did not have firm intentions until election day, or that they switched preferences. A favored explanation for those who conducted FSLN-leading polls is that voters changed their minds during the final days of the campaign, when polling was forbidden by law. There are theories, for example, that argue that Daniel Ortega's failure to address squarely the economic crisis facing the country and, particularly, his failure in the last days of the campaign to halt the military draft for the war against the Contras led previous supporters to vote for UNO in desperation. While there may have been a large shift of voters from the FSLN camp to UNO at the end of the campaign, this explanation requires one to discount entirely the many UNO-leading polls and the fact that polls on both sides showed voters reporting that their intentions were "firm."

Assuming for the moment that "real" vote intentions did not change massively at the end of the campaign, response error appears to offer the best explanation for the disparity among the polls. What role did voter screens, the form of the vote intention question, and the context in which it was placed play in the response error process?

The answer appears to be that these factors do not explain the difference between the FSLN- and UNO-leading polls. From the limited sample of questionnaires collected by the Commission on Nicaraguan Preelection Polls, it does seem to be the case that UNO-leading polls put the vote intention question deeper into the questionnaire, as a general rule. Some of the preceding questions in these instruments, concerning opinions about issues in the election campaign, could have primed respondents in a UNO direction. On the other hand, FSLN-leading polls commonly interposed some questions between the voter screen and the vote intention item (though it might be argued that they were typically more innocuous questions).[13] Further, UNO-leading poll outcomes were not dependent on the vote intention question being placed after sensitive attitude items, if Borge's description of his national, February poll is accurate. He reported in an interview with Janice Ballou, a commission member, that the poll consisted only of questions on age, sex, and vote preference, but its findings were clearly in the UNO-leading camp (Ballou 1990).

Use of a secret ballot does not appear to explain the difference between UNO- and FSLN-leading polls, either. Only the Bendixen-Schroth polls fielded by LOGOS and the Belden and Russonello poll fielded by Itztani used the secret ballot consistently. The LOGOS efforts produced a dead heat in the October poll and an FSLN lead in January; Belden and Russonello showed an FSLN lead. Others used "secret ballots" if respondents refused to answer direct questions (e.g., Itztani in polls for Hemisphere Initiatives), or only in some of their polls (Borge reports using the technique in the early February poll). Thus, both FSLN- and UNO-leading polls employed the method in February, shortly before the election. While the precise secret ballot technique may have differed among the polls (some using clear plastic bags for holding ballots and others using wooden boxes), that difference also seems to distribute across polls with disparate findings.

---

[13] It is not entirely clear whether putting the vote intention question first in the questionnaire or preceding it with issue items is better. In an interview with the Commission on Nicaraguan Preelection Polls on February 2, 1990, FSLN government official Paul Oquist criticized the CID-Gallup October poll for using question order to bias vote-intention responses. Crespi (1988) notes that the vast majority of U.S. pollsters avoid placing issue items prior to the vote intention item. But Carlos Denton of CID-Gallup (at the AAPOR annual meeting in May, 1990) defended the CID practice as necessary for establishing rapport with respondents. Denton's view on the value of issue items before the vote intention question is shared by some U.S. pollsters, notably Jeffrey Alderman of ABC, quoted by Crespi (1988). Nancy Belden points out that, for CID-Gallup respondents, rapport may have meant that interviewers were open to hearing criticism of the government.

The use of different types of voter screens also does not appear to explain the difference between the UNO- and FSLN-leading polls. Voter registration and turnout were very high by U.S. standards (nearly 90% of eligibles in both cases). Both sides had polls with the "likely voter" percentage in the high 80s, close to the eventual turnout in the election, suggesting that a major source of preelection poll variance in the United States was not operating in this case.

To summarize the review of data collection techniques, there were, of course, differences among the various polling organizations, but they do not seem to be systematically different between the two major poll "camps." The most systematic difference appears to be that the UNO-leading polls tended to precede the vote intention question with more potentially priming questions than did the FSLN-leading polls. Even here the difference is not consistent.

Ironically, based on available information, one's overall view of the quality of the sampling, data collection, and professional conduct of the polling organizations would tend to favor the polls that departed most from the election results. The FSLN-leading polls were more likely to employ national samples of voting districts (JRVs) after the October voter registration. They used secret ballots for measuring vote intention more often and tended to place the vote intention question earlier in their questionnaires, where the intention item would not be affected by previous answers. Unlike the UNO-leading polls conducted by VIA CIVICA, they did not have volunteer party activists doing interviewing.[14]

Also important was the general tendency of those who produced FSLN-leading polls to open their operations to public scrutiny. We would have little idea about what went on in the Nicaraguan preelection polls were it not for the cooperation of firms and poll sponsors such as ECO, Itztani, LOGOS, Belden and Russonello, Central American University, Hemisphere Initiatives, VIA CIVICA (a UNO-leading poll exception) and the *Washington Post*. The significant questions that remain are still outstanding largely because firms and sponsors such as DOXA, Borge y Asociates, CID-Gallup, Mora y Araujo, Noguera y Asociates, *La Prensa*, and IPCE refused to share methodological information or to allow observation during the preelection period. While willingness to disclose methodology does not, obviously, insure proper technique or accuracy, the fact that those who produced FSLN-leading results generally behaved in a more professionally responsible way is certainly disconcerting and adds to the irony of the mismatch between their polls and the election results.

---

[14] VIA CIVICA received funding from the National Endowment for Democracy through a U.S. foundation, the International Foundation for Electoral Systems (IFES). Richard Soudriette, director of the foundation, reported to me as this review was being written that the VIA CIVICA polls were conducted by VIA CIVICA volunteers, without technical assistance either from IFES or from any independent polling firm (see n. 17).

THE PEN EXPERIMENT

I have been unable to identify any systematic methodological differences that would explain the disparity between the UNO- and FSLN-leading polls. This observation leads back to earlier comments about the meaning and context of "vote intention" in the 1990 Nicaraguan election. As I have noted, there were a variety of reasons why measuring this concept was difficult. The unprecedented nature of the election, the novelty of polls, the extreme political polarization that called the idea of neutral measurement into question, the various sorts of intimidation during the campaign period (including some ersatz "polling" efforts), and the admonitions from the Supreme Electoral Council that "votes are secret," all contributed to an environment in which respondents could reasonably say one thing in a poll and do something else in the voting booth.

These are clearly post-hoc observations. Of course, differences between attitudes and behavior, between vote intentions and actual votes, are to be expected in every election, and those who conducted polls in Nicaragua were not unmindful of the likely disjunction. The use of secret ballots by some organizations is an example of this awareness. The experience of some pollsters in Latin American elections led them to believe that they could get accurate readings of the Nicaraguan electorate despite the difficulties. For example, Greenberg-Lake's press releases noted their success in predicting election outcomes in El Salvador and Paraguay, where the election circumstances were held to be analogous. Further, the volubility of Nicaraguan respondents in a variety of polls, their apparent willingness to criticize the Sandinista regime, and their generally reported belief that the election would be conducted fairly reinforced the conviction that vote intention measures would be reasonably accurate.

However, the level and nature of the disparity between election results and many of the polls was remarkable. It is clear, in retrospect, that respondents were saying one thing to one group of polling organizations and the opposite to another group. Schuman's "pen experiment" (1990), an offshoot of his participation on the Commission on Nicaraguan Preelection Polls, provided a timely empirical explanation for the respondents' behavior. His research showed that respondents would report markedly different vote intentions depending on their perception of interviewers' political affiliations.

The experiment flowed from the commission's first visit to Managua in early February. On that occasion, Schuman and I were told by UNO supporters that polls showing a Sandinista lead were not to be trusted, since "people were afraid" to give their actual opinions. Schuman contracted with ECO, the firm that was most forthcoming in response to our questions, to perform an experiment that would manipulate the apparent political affiliations of interviewers to see the impact on reported vote intentions. Raul Obregon of ECO was most anxious to do the study because, he related, he wanted to gather evidence concerning a legend about Nicaraguans' responses to changes in political regimes. The legend says that, during the civil wars over the last 150 years, villagers would keep handy the portraits of rival generals who were vying for power. The appropriate portraits could be displayed when the various contending armies would march through town. Obregon wanted to know if the same sort of variable show of allegiance would take place in polls, if they were perceived to be partisan.

The experiment was conducted during the first two weeks of February. The findings, reported elsewhere in more detail,[15] not only showed that respondents tended to give vote intentions that matched the apparent political leanings of the interviewers, but that they "assigned" interviewers an FSLN political leaning if none was readily apparent. Interviewers were directed to do interviews in each of three conditions—using a Sandinista logo pen, a UNO logo pen, and a "neutral" pen. Respondents, randomly assigned to the conditions, reported large vote intention majorities for the FSLN in the Sandinista and "neutral" pen condition, and a 56-44 split for UNO in the UNO pen condition.

The striking differences among the experimental treatments replicate, in microcosm, the split between the UNO- and FSLN-leading polling camps. The findings suggest that, if respondents were given clues about the political leanings of the polling firm, they would tend to report voting intentions in line with their perception. If clues were absent, respondents apparently assumed that the interviewer represented the government and gave responses supporting the FSLN. The fact that the pen cues were very subtle suggests how powerful the perceived political affiliation effect could be.

Perceived political affiliation appears to offer a major explanation for the disparate preelection poll results. It helps to account for the wide, consistent difference between the UNO- and FSLN-leading polls, regardless of the methods used to conduct them or their proximity to election day. It suggests that polling firms did not have to do anything untoward in order for their findings to be slanted in one direction or the other; in fact, it suggests that there was little they *could* do to counter the effect of perceived affiliation. "Neutrality" was not a solution.[16]

---

[15] Schuman reported the findings of the experiment in an Op Ed piece (1990) published shortly after the election.

While the "pen experiment" provided a crucial insight into the workings of the preelection polls, it, too, leaves some unanswered questions. First, it is not clear what motivated respondents to match their voting intentions to the perceived political affiliation of the polling firm. Fear was the explanation suggested by UNO partisans. If we believe that respondents were afraid to report UNO voting intentions, then they tended to give perceived FSLN-affiliated polls safe, untruthful responses and told the truth when queried by someone perceived to be affiliated with the opposition. Partisan activists from VIA CIVICA would get more truthful responses than neutral professional interviewers, according to this view. On the other hand, other forms of self-presentation may have been at work. For example, respondents may have tended to match the perceived affiliation simply out of deference or in line with the *gueguense* tradition (see n. 18 below). Since a number of the FSLN-leading polls were sponsored by U.S. organizations, respondents may have tended to give support to the Sandinistas out of a sense that they should show a united front against the U.S. pressure that had put the country in dire straits. In any of these cases, vote intentions expressed to neither side were "true," and the accuracy of the UNO-leading polls was coincidental.

Another set of unanswered questions concerns how respondents determined the political affiliation of the poll takers. We know that it was straightforward in some cases, as when interviewers for Borge y Asociates introduced themselves as doing polling for *La Prensa,* the Chamorro-family newspaper. It may also have been easy for respondents to decide that polls conducted by ECO were FSLN affiliated, because interviewers told respondents that the sponsor was Central American University, a Jesuit school that received government subsidies. But could respondents throughout the country have known that Itztani was directed by a man who was once affiliated with the Sandinistas, and would it have mattered if they could? How could they surmise that LOGOS was FSLN leaning when interviewers said nothing about the sponsor of the poll and the firm's director had assiduously avoided publicity? How could they decide that the Argentine firm, Mora y Araujo, Noguera y Asociates was UNO affiliated? If respondents tailored their responses to correspond with the perceived affiliation of the poll takers, it is important that we understand how they came to their perceptions. The partisan press may have played some role in publicizing affiliations for polling organizations, but it could not have done so in many cases. The "pen experiment" shows that perceived affiliation makes a difference in vote intention reporting, but it does not show why or how.[17]

---

[16] The finding that respondents in the "neutral" pen condition in the experiment tended to give vote intentions like the FSLN condition suggests that "neutral" poll takers were viewed as probable government representatives or sympathizers. This is plausible, as we know from Paul Oquist (1990) that the Sandinistas did unreported polling during the campaign, using government activists as interviewers. On the other hand, we need to consider the possibility that the "neutral" condition would take on a different perceived affiliation if the fieldwork firm were a known UNO affiliate—e.g., Borge y Asociates. In Schuman's experiment, the interviewers were identified as coming from Central American University, an entity that may have been perceived as Sandinista affiliated.

SUMMARY

As promised at the beginning, this review has not uncovered the "truth" of the Nicaraguan preelection polls. I have mainly discounted some explanations for the poll disparity. Although there remain holes in the argument that cannot be filled in now, the answer to the Nicaraguan poll quandary appears to lie in a real or believed partiality of the polls. The main question for respondents when confronting pollsters was, "Which side are you on?" Then, because of fear, deference, or the *gueguense* tradition of dissembling before authority,[18] respondents tended to tell interviewers what they thought they wanted to hear.

---

[17] Differences among the polling organizations in the composition of their interviewer and supervisor corps may shed a little light on how respondents identified the rival "camps," though information is sketchy. The following information is gleaned from the pollsters' press releases or reports. On the UNO-leading side, CID-Gallup, in its October poll, employed 25 interviewers, some Costa Rican and some Nicaraguan, with at least high school education and "from every political party." They carried CID identification. Mora y Araujo, Noguera y Asociates employed only 10 interviewers "with no political or research institute affiliation." Gender of interviewers is unknown for either firm. (It has been said that young women were more apt to be identified as Sandinistas.) DOXA interviewers were educated, professional people recruited through an advertisement in *La Prensa*, the UNO partisan newspaper, according to comments on this article by Nancy Belden, quoting Gustavo Mendez, Director of DOXA. A VIA CIVICA report on its January poll states that interviews were conducted by 300 teachers and students who gave their support to VIA CIVICA. Their training was done by VIA CIVICA's Training Department. Their instructions say that they should clearly state their affiliation and, if necessary, explain the VIA CIVICA mission.

On the FSLN-leading side, LOGOS reported employing 24 interviewers, mainly women, and ECO reported having 148 interviewers, "the vast majority" women. LOGOS insisted upon some college education, and ECO at least high school. Both firms claimed to screen applicants for impartiality, objectivity, and ability to get "in the door." ECO interviewers carried ECO and University of Central America identification. Nancy Belden, whose firm, Belden & Russonello, conducted the *Washington Post* poll, reported in comments on this review that the Itztani interviewer pool for that poll was largely women in their 20s and 30s, some of whom had worked on polls for the Morales Foundation. Finally, I heard other reports that interviewers sometimes worked for multiple organizations. ECO insisted, as a condition of employment, that they provide services exclusively.

[18] Martin Saballos, director of LOGOS, wrote a memo (1990) after his second poll for

While, in retrospect, the UNO-leading surveys appear to have been more accurate than the FSLN-leading studies, it is not clear that respondents were more truthful in the former polls. They may have felt free to open up to UNO-affiliated interviewers, or they may simply have given what they thought would be a pleasing response, without making a connection to their eventual "secret" ballot. We will probably never know whether the UNO-leading polls were actually or coincidentally more accurate than the others.

The Nicaraguan Polls Commission: History and Lessons

As I have noted, most of what is known about the Nicaraguan preelection polls is the result of the work of a commission formed to study them, funded by the Ford Foundation in Mexico. While this group has become the major source of knowledge about the polls, it also became embroiled in the partisan warfare of the Nicaraguan election. Since it was refused access to information by one of the polling "camps," it is important to understand various aspects of its history.

The commission was one of numerous unofficial groups observing various aspects of the Nicaraguan election. It was motivated by the same desire as that of other groups: to discover and publicize events and developments that might affect either this extraordinarily important election's outcome or its perceived legitimacy. In the case of preelection polls, scrutiny was deemed necessary because the major factions could claim, through polls, that they had been robbed if the election turned out against them. And while the United Nations and others were observing and documenting various aspects of the election process, no one was observing polls until the commission was formed.

The Nicaraguan poll observation effort was the brainchild of Ignacio Martin-Baro, the Salvadoran Jesuit public opinion scholar who was murdered in 1989, apparently by the Salvadoran army, along with several of his colleagues at the Jesuit university in San Salvador. Martin-Baro, a University of Chicago Ph.D., was instrumental in the founding of PROCADOP, the Central American Public Opinion Program, a group of survey research programs in the region that has carried out joint research on political culture and has promulgated a set of disclosure standards. The Nicaraguan commission, in turn, was a product of PROCADOP discussions in which Martin-Baro participated before

---

Univision that cited the *gueguense* tradition, expressing reservations about the validity of poll data. The tradition, like the portrait legend cited by Raul Obregon, states that peasants put on "masks" to talk to colonial authorities (and laughed at the authorities from behind them).

he died.[19] Following his murder, several PROCADOP members sought to carry out Martin-Baro's idea for a poll monitoring group in Nicaragua. Daniel Lund, a historian at the Universidad Autonoma Metropolitana (Mexico), and William Bollinger, director of the Interamerican Research Center in Los Angeles, were most instrumental in securing the funds for the project and in organizing the group of observers.

While a number of commission observers, including myself, had no previous connection with PROCADOP, the genesis of the observation effort with the Central American consortium led to what was probably an inevitable conflict with some pollsters in Nicaragua. PROCADOP had been critical of the conduct of CID-Gallup, the Costa Rican polling firm, in refusing to disclose its methodology in several surveys for the United States Information Agency (USIA) in the mid-1980s. These surveys, used by the Reagan administration to support its Latin American policy, were critiqued by Coleman and Sigelman in a poll review in *Public Opinion Quarterly*.[20] (The authors of that review, like the Nicaraguan Commission vis-à-vis the UNO-leading polls, also were in the position of having to rely on less data than desirable about the surveys' methodology, having received only partial information from USIA after a Freedom of Information request.) Coleman and Sigelman's review was preceded by an AAPOR session on the controversy, in which William Bollinger of PROCADOP participated. Bollinger and Daniel Lund also had previously aired the critique of the USIA surveys in *The Nation* (Bollinger and Lund 1988).

Bollinger's position as a prominent critic of U.S. government surveys in Central America and his central role in organizing the Nicaraguan observer group meant that the commission was viewed by pollsters in the UNO "camp" (including CID-Gallup) as having a political agenda, instead of a methodological interest. This was one reason why the commission received little or no cooperation from UNO-leading pollsters, despite repeated attempts at contact and despite the fact that a number of commission members had not been party to the earlier USIA dispute.

Victor Borge argued, at the 1990 AAPOR meeting, that U.S. academics had been manipulated by Bollinger to support the Sandinistas. *La Prensa* had branded the commission as a Sandinista front organization during the first observation visit in early February. Like interviewers for the various polling organizations in Nicaragua, commission members seemingly had to have a "side," and PROCADOP's history of criticism of U.S.-sponsored polling efforts enabled actors in Nicaragua to assign one to commission members, in spite of earnest attempts to stick to the methodological issues and remain above the partisan fray.

---

[19] Martin-Baro was honored posthumously at the 1990 AAPOR meeting for his remarkable polling work in El Salvador, and for his efforts to further survey research in Central America.

[20] See Coleman and Sigelman 1988. William Millard, a former USIA project officer, responded to Coleman and Sigelman's critique in the Spring, 1989, issue of *Public Opinion Quarterly* (53:134–35).

The commission's perceived partisanship was not the only reason (or even, perhaps, the most important reason) why some UNO-leading pollsters denied information to the observer group. Borge claimed that his methods were proprietary, and he did not want to give up his perceived commercial advantage. Carlos Denton of CID-Gallup objected to outsiders trying to impose a set of disclosure standards on his organization (he claimed to adhere to the code of the European Society of Marketing and Attitude Research).[21]

In spite of the noncooperation, the commission did manage to gather a considerable amount of material on the Nicaraguan polls, and its work continues in preparation of a book. Some U.S. members of the commission, notably Janice Ballou, are now seeking to establish closer ties with Central American pollsters so that cross-cultural methodological issues can be explored more easily and systematically. The documentation of the Nicaraguan polling matter (a resource for future scholars) and expanded interest in public opinion research in Central America suggest that the benefits of the commission's work outweigh the costs of the controversy.

The burgeoning growth of democracy around the world will give the survey research community many chances to observe polling in countries with little free election experience. Comparisons can be drawn between the Nicaraguan experience and that of other emerging democracies. We can hope that a spirit of professionalism and scientific inquiry will support future investigations, so that mysteries such as that posed by the Nicaraguan polls of 1990 will have a better chance of being solved.

References

Ballou, Janice. 1990. March 20. Interview with Victor Borge.

Bollinger, William, and Daniel Lund. 1988. "Mixing Polls and Propaganda." *The Nation* 246:635–638.

---

[21] The ESOMAR International Code of Marketing and Social Research Practice and the AAPOR Code of Professional Ethics and Practices both address the issue of disclosure. The ESOMAR code, however, appears to be somewhat less strict than AAPOR's, stating that researchers are "entitled" to release information on how a study was conducted if the findings are made public. The AAPOR code "imposes the obligation" upon researchers to report or make available information about how the research was conducted when findings are made public.

Coleman, Kenneth, and Lee Sigelman. 1988. "Poll Review: 1985 USIA Central American Surveys." *Public Opinion Quarterly* 52:552–56.

Crespi, Irving. 1988. *Pre-election Polling: Sources of Accuracy and Error*. New York: Russell Sage Foundation.

La Prensa. 1990. "Gran Sabotaje en el equipo de *La Prensa*." February 3.

Oquist, Paul. 1990. February 2 Interview with the Commission on Nicaraguan Preelection Polls.

Saballos, Marvin. 1990. "Social Psychology of Nicaraguans and Public Opinion Polls." Unpublished Memorandum. Centro de Consultoria y Capacitacion de la Communication.

Schuman, Howard. 1990. "3 Different Pens Help Tell the Story." *New York Times*. March 7.

A0610-1

Public Opinion Quarterly

BEST AVAILABLE COPY

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR Advisory Committee

Eleanor Singer, Chair (1987–91)
Herbert Abelson (1988–92) Gladys Engel Lang (1989–93)
W. Russell Neuman (1990–94)

Stanley Presser and Kathleen Frankovic,
AAPOR Councillors-at-Large

Howard Schuman, *Editor*

Office Manager Carlotta Nosel
Assistant to the Editor Katherine Bischoping

The Polls
Reviews: Stanley Presser, Seymour Sudman
Trends: Robert Y. Shapiro

EDITORIAL BOARD

Herbert I. Abelson
Robert P. Abelson
Sandra J. Ball-Rokeach
Larry M. Bartels
Nathaniel Beck
Sandra H. Berry
Lawrence D. Bobo
Pamela Conover
Jean M. Converse
Stanley Feldman
Richard J. Fox
Deborah R. Hensler
Shanto Iyengar
Graham Kalton
Gary King
Jon A. Krosnick
Richard A. Kulka
Kurt Lang
McKee J. McClendon
William J. McGuire
Michael MacKuen
Jack M. McLeod
Elizabeth Martin
John Mueller
Harold Nisselson
Helmut Norpoth
Vincent Price
Willard L. Rodgers
Norbert Schwarz
Robert Y. Shapiro
Lee Sigelman
Tom W. Smith
John Sullivan
John Zaller

Public Opinion Quarterly (ISSN 0033-362X) Volume 55, 4 issues (1991)

© 1991 by the American Association for Public Opinion Research.

Published quarterly: Spring, Summer, Autumn, and Winter, by the University of Chicago Press, 5720 S. Woodlawn Ave., Chicago, IL 60637. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $40.00; individuals 1 year $20.00. AAPOR member subscription rate, U.S.A.: individuals 1 year $18.00. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $18.00. In Canada, add 7% GST to subscription price. Outside of U.S.A., add $3.00 for each year's subscription to cover postage. Japanese subscription agent: Kinokuniya Company Ltd.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to: Journals Fulfillment Department, University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Institute for Social Research
P.O. Box 1248, Ann Arbor, MI 48106
Business, Subscriptions, Back Issues: University of Chicago Press,
P.O. Box 37005, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, NJ 08540

Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. ♾

PUBLIC OPINION QUARTERLY

VOLUME 55 / FALL 1991 / NUMBER 3

Articles

Race-of-Interviewer Effects in a Preelection Poll: Virginia 1989 313
STEVEN E. FINKEL, THOMAS M. GUTERBOCK, and MARIAN J. BORG

Response-Time Measurement in Survey Research: A Method for CATI and a New Look at Nonattitudes 331
JOHN N. BASSILI and JOSEPH F. FLETCHER

Work and Leisure: On the Reporting of Poll Results 347
RICHARD F. HAMILTON

Comment on "Work and Leisure: On the Reporting of Poll Results" by Richard F. Hamilton 357
F. THOMAS JUSTER and FRANK P. STAFFORD

Do Open-ended Questions Measure "Salient" Issues? 360
JOHN G. GEER

Voters' Intermediation Environments in the 1988 Presidential Contest 371
PAUL ALLEN BECK

Hardly Ever or Constantly? Group Comparisons Using Vague Quantifiers 395
NORA CATE SCHAEFFER

The Building of a Republican Regional Base in the South: The Elephant Crosses the Mason-Dixon Line 424
MARTIN P. WATTENBERG

Bringing Strong Ties Back In: Interpersonal Gateways to Political Information and Influence 432
BRUCE C. STRAITS

The Polls

Review: A Critique of The Kinsey Institute/Roper Organization National Sex Knowledge Survey 449
TOM W. SMITH

Intent and Purpose of The Kinsey Institute/Roper Organization National Sex Knowledge Survey: A Rejoinder 458
STEPHANIE A. SANDERS, MARY ZIEMBA-DAVIS, CRAIG A. HILL, and JUNE M. REINISCH

Book Reviews

Eric R. A. N. Smith, *The Unchanging American Voter* 463
STEPHEN EARL BENNETT

Jean-Noël Kapferer, *Rumors: Uses, Interpretations, and Images* 465
GARY ALAN FINE

Floyd J. Fowler, Jr., and Thomas W. Mangione, *Standardized Survey Interviewing: Minimizing Interviewer-related Error* 467
THERESA J. DEMAIO

Cedric Herring, *Splitting the Middle: Political Alienation, Acquiescence, and Activism among America's Middle Layers* 469
JOEL D. ABERBACH

AAPOR Proceedings

Conference Report 471

AAPOR Award 473

Presidential Address: *Trashing the Polls* 474
JOAN S. BLACK

Conference Program 482

Annual Membership Meeting 494

NOTICE TO CONTRIBUTORS

Prospective contributors should refer to the detailed Notice to Contributors that appears in the Spring 1991 issue of *POQ*.

RACE-OF-INTERVIEWER EFFECTS IN A PREELECTION POLL

VIRGINIA 1989

STEVEN E. FINKEL
THOMAS M. GUTERBOCK
MARIAN J. BORG

Abstract All published preelection surveys of the 1989 Virginia gubernatorial contest overestimated the vote share of the black candidate and eventual victor, Douglas Wilder. We offer a "social desirability" interpretation of the polls' inaccuracies and hypothesize that claiming support for Wilder was the socially desirable response for some whites, especially when the interviewer was black. We show a race-of-interviewer effect on the vote intention of white respondents of 8–11 percentage points in a preelection survey of Virginia voters. The effects were greatest among white Democrats and among whites who were more uncertain of their vote intention. We discuss the implications of these findings for race-of-interviewer research and for improving the accuracy of preelection forecasts in contests with black and white opposing candidates.

The vote returns that came in on the evening of the 1989 gubernatorial election in Virginia held an unexpected surprise. All the published preelection polls, as well as an exit poll completed on election day and broadcast as the polls closed, showed the black candidate, Democratic Lieutenant Governor Douglas Wilder, as having a decided lead over his white opponent, Republican Marshall Coleman. Estimates of Wil-

STEVEN E. FINKEL is associate professor of government and foreign affairs and associate director of the Center for Survey Research at the University of Virginia. THOMAS M. GUTERBOCK is associate professor of sociology and director of the Center for Survey Research at the University of Virginia. MARIAN J. BORG is research associate at the Center for Survey Research and a Ph.D. candidate in the department of sociology at the University of Virginia. An earlier version of this paper was presented at the annual meeting of the American Association for Public Opinion Research, Lancaster, Pa., May 1990. Funding for the collection of the data was provided by grant RF-87038 from the Equal Opportunity Division of the Rockefeller Foundation. We thank Larry Sabato for his advice and support of this Virginia poll project. We are indebted to Scott Keeter and the *POQ* reviewers for their valuable comments on earlier drafts.

Public Opinion Quarterly Volume 55:313–330 © 1991 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/91/5503-0002/$2.50

der's lead in the final preelection polls of the four major survey organizations in the state ranged from a 4 to an 11 percentage point advantage.[1] But the actual results gave Wilder a victory by the narrowest of margins, 50.3% to 49.7% (just 6,741 votes out of a total of 1.79 million ballots cast). The overestimation of Wilder's victory margin in the 1989 preelection polls was similar to the patterns seen in several other black-white statewide contests, such as the 1982 Bradley-Deukmejian gubernatorial election in California, and Wilder's own victory for lieutenant governor in Virginia in 1985. In each of these elections, the black candidate enjoyed a substantially greater share of the votes in preelection polls than he actually garnered on election day.

Why were the polls wrong in Virginia? The unrealistic results cannot be attributed solely to sampling error: several polls found the Wilder margin to be statistically significant, and the fact that all the polls fell on the same side of the true mean suggests a systematic bias rather than random error. No doubt some of the variability in the poll results has its origin in the "normal" culprits responsible for preelection polling inaccuracies, for example, imperfect sampling procedures, last minute shifts of voting preferences, faulty estimates of the likely electorate, and improper allocation of undecided voters (Bradburn and Sudman 1988; Crespi 1988; Roll and Cantril 1980). To this list, however, we add another source of preelection polling error arising specifically in elections with black and white opposing candidates: under certain contextual conditions, white respondents are more reluctant to report intentions to vote for the white candidate, and more willing to report intentions to vote for the black candidate, than they are to cast their ballots for those candidates on Election Day. The existence of this type of "social desirability bias" in preelection polls is difficult to prove since the institution of the secret ballot rightly prevents the researcher from checking any respondent's expressed voting intention against his or her actual vote. However, the existence of social desirability bias can be inferred from a number of more readily measurable patterns, the most significant of which is that white respondents tend to report different voting intentions depending on whether the person who interviews them is white or black. In this paper, we present evidence of a race-of-interviewer effect in a preelection survey of Virginia voters, explore the conditions under which the effect is most likely to

[1] The major polls and their final estimates of voter preference were: Mason-Dixon Opinion Research, Inc. (October 30–November 1), Wilder 48 percent, Coleman 44 percent, undecided 8 percent ($N = 825$); *Washington Post* (October 30–November 2), Wilder 52 percent, Coleman 41 percent, undecided 7 percent ($N = 751$); *Richmond Times-Dispatch* (October 27–November 1), Wilder 45 percent, Coleman 36 percent, undecided 19 percent ($N = 852$); and the Commonwealth Poll, Virginia Commonwealth University (September 21–October 19), Wilder 44 percent, Coleman 38 percent, undecided 18 percent ($N = 451$). All were conducted by telephone.

occur, and assess the implications of these effects for electoral and other opinion surveys dealing with racial issues.

Prior Research

Studies reporting effects of the race of the interviewer on responses to survey questions are not uncommon in public opinion research. Several studies show tendencies of black respondents to report more favorable attitudes toward whites to white than to black interviewers (Anderson, Silver, and Abramson 1988b; Schuman and Converse 1971), and more recent work has shown significant race-of-interviewer effects on white responses to racially oriented questions as well (Campbell 1981; Cotter, Cohen, and Coulter 1982; Hatchett and Schuman 1975–76; Sudman and Bradburn 1974). While most research has found effects mainly on race-related items, Schuman and Hatchett (1974) also report significant race-of-interviewer effects on black attitudes about the political system, patriotism, and democratic values. The most common explanation for all these effects is that individuals seek to avoid offending interviewers of the opposite race because of "interpersonal deference," "courtesy to a polite stranger," or more simply a general desire to be agreeable. On certain kinds of survey items, individuals react in part to the social pressure of the interview situation and tend to respond based on their expectations of the interviewers' preference. Much related work suggests that individuals tend to give socially desirable responses to other survey questions, for example, those relating to voter turnout (Anderson, Silver, and Abramson 1988a; Silver, Anderson and Abramson 1986); the race-of-interviewer effect shows that, especially with racially oriented questions, the "socially desirable" response depends in part on who is asking the question.

Such logic could certainly apply to expressed voter preferences in certain preelection surveys as well. Pollsters typically expect some respondents in preelection surveys to refuse to state their vote intentions and others to misstate it, yet they generally have no reason to suppose that the probabilities for refusal, concealment, and prevarication are systematically different among supporters of either party's candidate. While this tacit assumption of equiprobability of error may hold in most elections, in election contests with black and white opposing candidates, social desirability pressures of various forms may be present and may significantly bias estimates of the likely vote outcome.

In general, individuals may feel some social pressure to claim support for a black candidate in the interview situation. Few individuals in modern American society want to appear antiblack, as social norms regarding race have become increasingly egalitarian in recent years (Schuman, Steeh, and Bobo 1985). Voicing support for a white candidate in a black-white contest, however, might be interpreted by the interviewer as a race-based vote; hence, there may be a subtle pressure on white voters to claim to support the black candidate, no matter what their true preference (and actual future behavior) may be.

This pressure should be felt most strongly, and be more directly observable, under a specific contextual condition, that is, when a white respondent is being surveyed by a black interviewer. In many recent local and statewide elections pitting white and black candidates against one another, blacks have been nearly unanimous in their electoral support of the black candidate. Whites have varied in their degree of bloc voting, with varying proportions of whites giving their vote to the black candidate in different contests. Thus, respondents who are asked their voting preference will have a good idea of the likely preference of a black interviewer but will not be as certain about the likely preference of a white interviewer. In the eyes of the respondent, a white interviewer might be quite likely to support the white candidate, but a black interviewer would be viewed as almost certainly supporting the black candidate. While individuals surveyed in other kinds of electoral contexts could also make estimates of the preferences of interviewers based on gender, accent, or race as well, the special circumstance of black and white opposing candidates may make both interviewer preferences and racial divisions in the electorate more salient for the respondent. Thus, to the extent that social desirability motivates individual responses, race-of-interviewer effects should be seen most clearly in surveys during these types of election contests.

The 1989 Virginia election in particular is an ideal setting in which to test for such social desirability effects. Not only was Douglas Wilder bidding to become the nation's first elected black governor, but he was doing so with much media fanfare in the heart of the old Confederacy. The Democrats had won the two previous gubernatorial elections in the state (Charles Robb in 1981 and Gerald Baliles in 1985), and the incumbent Baliles enjoyed widespread popularity, making Wilder a competitive candidate at the outset of the contest. Wilder ran a highly credible campaign as a qualified, major-party candidate, ably portraying his white opponent as tending toward conservative extremism. Coleman's attacks on various episodes in Wilder's earlier career failed to do much damage to Wilder's public credibility. Thus, white voters who had supported earlier Democratic candidates had few convenient "excuses" for withholding support for Wilder, no obvious reason for it other than a negative reaction to the candidate's race.

By the end of the contest, moreover, it was apparent that the race issue had been a significant part of the overall electoral landscape. The candidates themselves rarely, if ever, broached the racial dimensions of the contest during the campaign, but the national media "insisted on branding Wilder a black politician [even though] he was determined not to run as one" (Sabato 1991, chap. 5). And while white voters themselves denied that race would play a role in their decisions, with only 12 percent saying in our survey that the race of the candidates would be "very" or "fairly important" in their vote decision, 64 percent of whites also perceived Wilder as likely to help blacks "a lot" once in office compared to only 16 percent for Coleman, indicating the potential for racially polarized voting. The election outcome, moreover, showed that Wilder ran 4 percentage points behind the victor in the lieutenant governor's race, Donald Beyer, and 13 points behind Attorney General Mary Sue Terry, both of whom are white. These results fueled much speculation that race may have been an important, though "subterranean and unmentionable" aspect of the electoral campaign (Sabato 1991, chap. 5). Given this context, we hypothesize that social pressures related to race would influence responses to preelection surveys: voicing support for Wilder would be the socially desirable response for whites, especially when the interviewer was black.

Data

The data used in the study are from a statewide survey of Virginia residents conducted by the University of Virginia's Center for Survey Research between October 22 and November 2, 1989. A total of 362 respondents were surveyed, of whom 256 were registered voters. The survey was conducted by telephone, with approximately one-third of all completed interviews conducted by black interviewers. The interviewers were graduate and undergraduate students at the University, many of whom had previous telephone interviewing experience. All were carefully trained in techniques for telephone interviewing.

Although there is no face-to-face contact between interviewers and respondents with telephone interviews, several studies have reported race-of-interviewer effects in telephone surveys (Cotter, Cohen, and Coulter 1982; Meislin 1987; Morin 1989). Three processes may lead to these effects: respondents may correctly identify the race of the interviewer through accent or speech pattern and adjust their answers accordingly; interviewers state their names at the outset of the survey, and names can sometimes be taken as cues to racial identity; or the interviewer may (consciously or unconsciously) use verbal cues, into-nations, or inflections that elicit certain responses to questions of particular importance or salience to the interviewer.[2] Whatever the mechanism, we may expect a race-of-interviewer effect in surveys that deal with racial issues or that are conducted in racially polarized electoral environments, even when the interview is conducted by telephone.

Telephone numbers were fed to each interviewer station randomly from a larger pool of randomly selected telephone numbers of Virginia households. We limit our analysis to only white and black interviewers, and report no results for the 59 respondents interviewed by Asian-American or Asian students. Because of limitations of sample size, we focus only on white respondents, 149 of whom were surveyed by white and 103 by black interviewers.[3]

The voting preference of each registered voter was determined by asking: "If the election for Governor were held today, who would you vote for? The Democratic candidate, Douglas Wilder, or the Republican candidate, Marshall Coleman, or aren't you sure how you will vote?" For those who said they were "undecided," a follow-up question was asked: "Which way do you lean as of today, toward Douglas Wilder or toward Marshall Coleman?"[4]

Results

The results of the initial race-of-interviewer test are shown in table 1. As can be seen, reported preference for Wilder differs by 8 percentage points depending on the race of the interviewer (44 percent among white interviewers, 52 percent among blacks), while reported preference for Coleman differs by 11 points (44 percent among white interviewers, 33 percent among blacks). The percentages of undecided voters are almost equal, with a slightly higher figure among respondents

---

[2] Meislin (1987) reports the results of an earlier study in which over three-quarters of respondents correctly identified the race of the interviewer.

[3] The percentage of interviewed blacks in our sample was 16 percent, almost exactly their numbers in the Virginia electorate, but too small to permit reliable statistical analyses. Among these respondents, we did find that blacks interviewed by blacks were 11 percent more likely to claim support for Wilder than blacks interviewed by whites. This finding is consistent with the results shown below for white respondents, but we do not elaborate on the results for black respondents because of the very small number of cases.

[4] We should note that the sequence of questions may also act as a contextual factor that may heighten the race-of-interviewer effect. In our survey, vote intention was asked about two-thirds of the way through a 20-minute interview script. Party identification was measured well after the vote intention questions, but racial attitude questions were near the beginning of the interview, as were a series of questions asking respondents to identify the race and sex of various candidates. These questions may have increased the salience of racial orientations for respondents, but we doubt that they account fully for the patterns found in our survey.

Table 1. Voter Preference among Whites by Race of Interviewer

	White Interviewers (%)	Black Interviewers (%)
Wilder	43.8	52.2
Undecided	12.4	14.9
Coleman	43.8	32.8
N	105	67

NOTE.—$\chi^2 = 2.06$, 2 *df*, $p = .36$, Cramer's $V = .11$.

interviewed by blacks. The Wilder-Coleman percentage split differs rather markedly depending on the race of the interviewer; nevertheless, the differences in this bivariate table are statistically insignificant.

Part of the reason for the lack of a statistically significant effect is no doubt due to the small number of respondents in the sample. However, it was also the case that in our study, blacks were more likely to interview white Republicans than were white interviewers, as 46 percent of all completed interviews by blacks were with Republicans, compared to only 33 percent of completed interviews by whites.[5] Thus, the pool of white respondents interviewed by blacks was more likely to be pro-Coleman than respondents interviewed by whites, mitigating the effect of the race of the interviewer in the bivariate table. Table 2 shows the voter preference results for each party group separately, and here significant and more substantial race-of-interviewer effects can be seen.

For example, among white Republicans (including Independents who lean Republican), the percentage who favor Coleman differs by 18 points depending on the race of the interviewer (69–51 percent).

---

[5] The assignment of telephone numbers to interviewers was done completely randomly, and scheduled callbacks were, whenever practical, assigned to the same interviewer. The differences in party affiliation among respondents interviewed by blacks and whites, then, seems to be due to more than chance (and the chi-square value for this percentage difference is 3.41, $p = .06$). Quite possibly, this pattern may be due to selective refusal rates among whites when surveyed by blacks over the telephone. Our data indicate that approximately 31 percent of all contacts that black interviewers made with all respondents resulted in an immediate refusal, while only 16 percent of contacts made by whites refused to be interviewed. While we do not know the party identification or other characteristics of the individuals who refused to be interviewed, we do know that respondents interviewed by blacks were disproportionately located in northern Virginia and other suburban locales throughout the state, areas with relatively strong Republican loyalties, as judged by previous statewide surveys. Thus, we infer that rural, white Virginians had higher refusal rates when called by a black interviewer. Differential refusal rates were probably heightened in our poll by our use of a respondent selection procedure at the outset of the interview as described by Troldahl and Carter (1964).

Table 2. Voter Preference among Whites by Race of Interviewer and Party Identification (in percent)

	Republicans		Independents		Democrats	
	White Interviewer	Black Interviewer	White Interviewer	Black Interviewer	White Interviewer	Black Interviewer
Wilder	26.5	27.0	16.7	63.6	70.5	94.7
Undecided	4.1	21.6	50.0	9.1	11.4	5.3
Coleman	69.4	51.4	33.3	27.3	18.2	0.0
N	49	37	12	11	44	19
χ^2	6.69		6.46		4.98	
	(.04)		(.04)		(.08)	
Cramer's *V*	.28		.53		.28	

NOTE.—Republican and Democratic groups include independent "leaners"; Independents include "apoliticals." Numbers in parentheses are level of probability.

The number of undecideds differs by 18 points, with support for Wilder equal at 27. The chi-square for the entire Republican subgroup is 6.69 with 2 *df*, significant at the .05 level. When interviewed by blacks, then, white Republicans are significantly less likely to report a preference for the white candidate and more likely to report being undecided. Among white Democrats (and Independents who lean Democratic), the results are similar in magnitude (Cramer's *V* of .28 for both party groups), but the Democrats show strong differences in their expressed support for both Wilder and Coleman depending on the race of the interviewer (percentage differences of 24 points for Wilder and 18 points for Coleman). This suggests that without taking into account potential race-of-interviewer effects, estimates of the likely vote for the black (Democratic) candidate may be biased particularly among white Democrats, while estimates of the likely vote for the white (Republican) candidate may be biased for both major party groups. Finally, among pure Independents, race of interviewer is related to vote choice as well. Although the sample size is extremely small for this group ($N = 23$), the table does indicate that a substantially greater percentage of respondents interviewed by blacks claimed support for Wilder than those interviewed by whites.

The results in this table suggest that bivariate comparisons of responses to white and black interviewers are not sufficient to determine a race-of-interviewer effect if the types of respondents interviewed by whites and blacks differ in systematic ways. These possibilities increase, of course, in surveys with smaller samples and if refusal rates to interviewers of different races differ based on some attitude or demographic factor relevant to the vote. Race-of-interviewer effects, then, can best be demonstrated after controlling for such factors in multivariate analysis. Table 3 shows the results of logistic regression analyses predicting individuals' probabilities of expressing a vote intention for Coleman and Wilder with race of interviewer and party identification as independent variables, along with other relevant political and demographic factors as control variables.

Two important political variables, aside from party identification, appear in the models: individual attitudes toward abortion, which was a major issue dividing Wilder and Coleman during the campaign, and racial attitudes, which potentially could exert independent influence on whites' votes for a black candidate in any black-white contest. We measured abortion "liberalism" by counting the negative responses to three questions concerning possible restrictions on a woman's ability to obtain a legal abortion.[6] We constructed the racial attitudes scale

---

[6] The exact wording of the questions was: (1) "The laws about abortion ought to be changed to make it more difficult for a woman to get an abortion"; (2) "It ought to be

Table 3. Logistic Regression Model Predicting Vote Intention for Coleman and Wilder

	Coleman Model		Wilder Model	
Variable	Coefficient	S.E.	Coefficient	S.E.
Race of interviewer (0 = black, 1 = white)	1.04*	.41	−1.00*	.42
Party identification	.61*	.12	−.60*	.12
Racial liberalism	−.38	.24	.41†	.22
Abortion liberalism	−.49*	.22	.83*	.25
Education	−.04	.08	.07	.08
Age	.01	.01	.00	.01
Income	−.17	.19	.34†	.20
Constant	−1.62		−1.17	
Percent correctly classified	72.7		77.3	
Mean of dependent variable	.40		.47	
Estimated R^2	.30		.34	
N	172		172	

Note.—Predicted probability of Coleman vote intention for "average" respondent and black interviewer = .24; predicted probability of Coleman vote intention for "average" respondent and white interviewer = .47; predicted probability of Wilder vote intention for "average" respondent and black interviewer = .64; and predicted probability of Wilder vote intention for "average" respondent and white interviewer = .40.

* Significant at .05 level (two-tailed).

† Significant at .10 level (two-tailed).

based on counting the racially "liberal" responses to four questions measuring support for racial equality, affirmative action programs, attitudes about the pace of improvement in societal conditions for blacks, and beliefs about the ease with which blacks and whites can get jobs.[7]

---

legal for a woman to get an abortion if she is pregnant as a result of rape or incest"; and (3) "If a young woman under the age of 18 wants to get an abortion, the law should require that her parents be notified even if she doesn't want them to know about it." There were four response categories for each question, ranging from strongly agree to strongly disagree. We took as the liberal response "agree" or "strongly agree" to question 2, and "disagree" and "strongly disagree" to questions 1 and 3.

7. The exact wording of these questions was: (1) "Black people and white people ought to be treated equally in every way"; (2) "These days it is a lot harder for a black person to get a good job in Virginia than it is for a white person"; (3) "Government and private companies should be making more effort to hire blacks and other minorities even if it means giving blacks preferential treatment"; and (4) "Some people are trying to improve conditions for blacks in Virginia. Do you think that those who want to improve conditions for blacks are trying to do it too quickly, not quickly enough, or just about right?" There were four response categories for each of the first three questions, ranging from strongly agree to strongly disagree. We took as the racially liberal response "agree" or "strongly agree" to the first three questions, and "not quickly enough" for the last.

We also included standard demographic variables such as education (measured in years of schooling), age, and income (measured as annual household income with five categories from "less than 10,000 dollars" to "greater than 60,000 dollars").

The results show that race of the interviewer is a significant factor in predicting expressed vote intentions for both Coleman and Wilder, independent of the effects of all other variables. While party identification, abortion, and, to a lesser extent, racial attitudes and income are all relevant to vote intention, whites also appear to respond to the cue represented by the interviewer's race in voicing their support for the white or black candidate. Because logistic models do not yield linear probability estimates, there is not a constant effect of the race of the interviewer on the probability of expressing an intention to vote for one or another candidate. However, one way of reporting the effect of the race of the interviewer is to show the differences in predicted probabilities of vote intentions for Wilder and Coleman among "average" white respondents who are interviewed by whites and by blacks, that is, among whites who are at the mean on all other independent variables (cf. Aldrich and Nelson 1984). The results of these calculations are shown in the note to table 3 and indicate that the average white respondent interviewed by whites has a .23 greater estimated probability of expressing intentions to vote for Coleman, and a .24 lesser probability of expressing intentions to vote for Wilder, than their counterparts interviewed by blacks. This is strong evidence for the existence and strength of a race-of-interviewer effect in this election survey.

Race-of-Interviewer Effects and Strength of Voter Commitment

Given that race-of-interviewer effects may occur in preelection surveys, it is of critical importance to discover the types of respondents among whom these effects are greatest, so that appropriate measures to improve the predictive capability of the poll may be taken. In previous race-of-interviewer research, some specification effects have been reported for education, social status, and place of residence, though the differences in the magnitude of effects for these demographic categories have been modest (Anderson, Silver, and Abramson 1988a; Campbell 1981; Hatchett and Schuman 1975–76). We tested for such differential effects by including an interaction term in the logistic regression model between the race of the interviewer and each of the other independent variables. Thus, for example, a race-of-interviewer and racial liberalism interaction effect would indicate that individuals who are more racially liberal were more (or less, if the interaction effect were negative) influenced by the interviewer's race than were racially conservative respondents.

In the interaction models predicting vote intention for Wilder, only the interaction of race of interviewer with party identification was statistically significant, indicating that Democrats were more likely to report differential vote intentions to whites and blacks than were Republicans. This finding confirms the pattern shown in table 2. While we cannot provide a definitive explanation for why this pattern occurs, one possibility is that white Democrats feel "cross-pressured" between their traditional party affiliation and race, and they minimize this conflict in the preelection interview by responding on the basis of the race of the interviewer to a greater extent than traditional Republicans.

In the model predicting vote intention for Coleman, none of the interaction terms with party, racial or abortion liberalism, or the demographic terms was statistically significant, and thus we conclude that the race-of-interviewer effect in general operated similarly across individuals with different political and attitudinal characteristics. However, we did test one additional interaction effect of special importance for preelection polling. A common hypothesis in studies that estimate question wording and order effects in surveys, though not one that has received clear empirical support (cf. Schuman and Presser 1981), is that when individuals are relatively uncertain of their preferences, they should be more likely to respond to "extraneous" contextual cues. In electoral polling, this suggests that individuals who are least committed to a given candidate may be more likely to be influenced by environmental cues, one of which is the race of the interviewer. The pattern reported above of white Democrats responding to the race of the interviewer more than other party groups could also be explained by this process, as Democrats may have less intense preferences because of the cross-pressure between party and race and, hence, more susceptibility to interviewer effects. Table 4 presents more direct evidence supporting the "uncertainty" hypothesis.

In this table, we divide the sample into two groups: those who expressed a vote intention on the initial vote preference question asked of all registered voters; and those respondents who claimed they were initially "undecided" and were asked the follow-up question concerning which candidate they "leaned" toward supporting. If the uncertainty hypothesis is correct, we would expect race-of-interviewer effects to be greater among those who say they are initially undecided than among those who offer a candidate preference with no further probe necessary. This hypothesis is strongly supported by the data presented in the table. For initially decided respondents, there is al-

Table 4. Preferences of Initially Decided Respondents and Probed Preferences for Initially Undecided Respondents by Race of Interviewer (in percent)

	Vote Preference of Initially Decided Respondents		Probed Preference for Initially Undecided Respondents	
	White Interviewer	Black Interviewer	White Interviewer	Black Interviewer
Wilder	55.1	58.7	22.2	38.1
Undecided	. . .	. . .	36.1	47.6
Coleman	44.9	41.3	41.7	14.3
N	69	46	36	21
χ^2	.04 (.85)		4.73 (.09)	
Cramer's *V*	.04		.29	

NOTE.—Numbers in parentheses are level of probability.

most no race-of-interviewer effect, with differences of 3 or 4 points in preference for a given candidate. Among whites who were asked the follow-up "leaning" question, however, race-of-interviewer effects are substantial and are statistically significant at the .10 level. Here reported vote intention for Wilder is 16 points greater for respondents interviewed by blacks than by whites, and vote intention for Coleman is 27 points greater for respondents interviewed by whites than by blacks.

In interpreting this result, we must consider an alternative explanation for the greater race-of-interviewer effect observed among those who initially claimed to be "undecided." It is conceivable that this is merely a selection effect: this would be so if the "undecided" response to the initial question was selected by some of our white respondents as a means of masking their vote intention. After all, white respondents who did not want to "offend" the black interviewer on the initial preference question could do so either by declaring their intention to vote for the black candidate or by reporting that they were undecided.

However, our data do not support this thesis. If white respondents were using the initial "undecided" response as a mask for their intentions, then we would expect more "undecided" responses on the initial vote preference question among those interviewed by blacks than by whites. But only 31.3 percent of the white registered voters interviewed by blacks initially claimed to be "undecided," no greater proportion than the 34.2 percent figure among those interviewed by whites. We can reasonably conclude that those reporting themselves to be "undecided" to black and white interviewers were equally genuine in the uncertainty of their voting preference.

We can assess the net impact of respondent uncertainty on the race-of-interviewer effect by expanding the logistic regression model for Coleman from table 3 to include dummy variables indicating the respondent's membership in the following categories: initially decided and interviewed by a white; initially decided and interviewed by a black; and initially undecided and interviewed by a white. The baseline category is all respondents who were initially undecided and were interviewed by blacks. We report the results from this model in table 5.

The table shows that the estimated difference in logit coefficients for initially decided individuals interviewed by whites and blacks is .61 (2.65–2.04), and this difference is statistically insignificant. Thus the estimated probability of expressing a preference for Coleman among initially decided voters does not depend on the race of the interviewer. For the initially undecided voters, however, the estimated difference in the logit coefficient for respondents interviewed by blacks and whites is 2.16, and this difference is statistically significant at the .01 level. The difference in effects corresponds to a .25 greater probability that the "average initially undecided white respondent" will express a vote intention for Coleman to a white interviewer than to a black, providing clear evidence that initially undecided voters are more likely to be influenced by the cues of the interview situation than voters with a stronger initial preference.

Discussion and Conclusions

We extend previous research by finding a significant race-of-interviewer effect on expressed voter preferences in a preelection survey, a result that, to our knowledge, has not previously been reported. The effect is, in electoral terms, relatively substantial, as whites are 8–11 percentage points more likely to voice support for the black candidate to black than to white interviewers, and to voice support for the white candidate to white interviewers than to black. This suggests that as such electoral contests become more frequent in American politics, pollsters and survey researchers need to be aware of the race-of-interviewer biases that may emerge on voter preference questions as well as on racial attitude and participation items.

Further, we find that the effects were concentrated among white voters who were least committed to a particular candidate, as mea-

Table 5. Logistic Regression Models Predicting Vote Intention for Coleman, including Uncertainty/Race-of-Interviewer Effect Interaction

Variable	Coefficient	S.E.
Initially decided, white interviewer	2.65*	.75
Initially decided, black interviewer	2.04*	.76
Initially undecided, white interviewer	2.16*	.77
Party identification	.63*	.12
Racial liberalism	−.48†	.25
Abortion liberalism	−.47*	.23
Education	−.05	.08
Age	.00	.01
Income	−.24	.20
Constant	−2.57	
Percent correctly classified	77.3	
Mean of dependent variable	.40	
Estimated R^2	.37	
N	172	

NOTE.—Predicted probability of Coleman vote intention for "average" initially undecided respondent and black interviewer = .05; and predicted probability of Coleman vote intention for "average" initially undecided respondent and white interviewer = .30.

* Significant at .05 level (two-tailed).

† Significant at .10 level (two-tailed).

sured by an "undecided" response on an initial vote preference question. Expressed support for the black candidate depended on the race of the interviewer most strongly among white Democrats as well. These patterns suggest that individuals are influenced by the race of the interviewer when they are unsure of their preferences, when they are asked to state their vote intentions when their opinions have not yet crystallized, or when their voting predispositions based on party and race are in conflict. If our findings can be generalized, they indicate that party identification and degree of certainty that an individual feels about his or her intended vote will be important specification factors in predicting the likelihood that individuals will be influenced in their responses by the race of the interviewer.

A more practical consequence of the race-of-interviewer effects seen here is that they may seriously distort preelection forecasting in black-white contests.[8] The extent of the bias depends, of course, on how large the effect is and the number of interviews with white respondents completed by blacks (as well as possible differences in white refusal rates when contacted by white and black interviewers). Determining whether responses to whites and blacks are the respondent's "true" attitude has proven difficult for much race-of-interviewer research, but in the electoral context we may use the actual aggregate vote distribution by race as a measure of accuracy. Judged by this standard, it is clear that the white interviewer data here provide a more reliable estimate of future vote intention. In our poll, support for the candidates was at 49 percent for Wilder and 39 percent for Coleman among all white respondents, and at 44 percent for each candidate among only those white respondents surveyed by white interviewers (see table 1). Given that Wilder's actual white support, as estimated by several methods, was approximately 40 percent on Election Day, it is clear that the white interviewer data yield more predictive accuracy.[9]

---

[8] At least one of the four major polling organizations in Virginia, the Commonwealth Poll from Virginia Commonwealth University, has confirmed a race-of-interviewer effect in their last preelection poll, although of a smaller magnitude than in our survey (Keeter 1990).

Our findings lead to one unambiguous recommendation: survey organizations must record the race of interviewers and check for these effects whenever they conduct polls in black-white electoral contests. In random telephone polls, it is scarcely possible to match the race of interviewer and respondent at the time of the interview. But in future black-white contests researchers may be well advised to use only the data obtained by white interviewers in making electoral predictions of the white vote, despite the loss of efficiency involved in doing so (cf. Fowler 1984, pp. 113–14).

Several qualifications, however, should be made to this recommendation. First, it is by no means inevitable that the racial context of all elections, or all survey situations in elections with black and white opposing candidates, will lead to race-of-interviewer effects. Thus we do not recommend unconditionally the use of only white interviewers to interview white respondents, even where this might be practical. Rather, we suggest using both black and white interviewers in order to test for race-of-interviewer effects in preelection surveys. If such effects are found, then the white interviewer data are likely to be more accurate.

Second, the finding that race-of-interviewer effects were found principally among undecided voters suggests that less bias will result if pollsters do not probe respondents for their preferences after an initial vote intention question. The drawback to this approach, of course, is that many more "undecided" voters will be present in published polls, making predictions of the ultimate vote outcome more difficult. But given the recent pattern of erroneous predictions by pollsters in black-white contests, this recommendation may be useful in forcing survey researchers to construct better means than presently used for allocating undecided voters in such racially oriented electoral contexts. One intriguing possibility is to allocate undecided voters by race (Guterbock and Finkel 1989; Hugick 1990; Sussman 1985), which, while not necessarily describing accurately the individual-level dynamics of preference formation during a campaign, nevertheless would have yielded an almost perfect prediction of the distribution of white votes in the 1989 contest.

---

[9] A CBS–New York Times exit poll published in *Time* (Shapiro 1989) showed white support at 39 percent. Further, given that most estimates of the composition of the Virginia electorate put the percentage of blacks at approximately 16–17 percent, the proportion of whites who supported Wilder—given close to unanimous Wilder support among blacks—could have been little more than 40 percent.

In general, we believe that the finding of a race-of-interviewer effect in a preelection survey should be used to alert the pollster to the possible hidden importance of race in the election. If race-of-interviewer effects exist, they indicate that some portion of the white electorate is responding to racial cues in the overall electoral context. Under some conditions, it may even be the case that white support for the black candidate that is reported to white interviewers in preelection polls is also overstated, especially among initially undecided voters. The interpretation of this pattern would be that respondents believe that the social norm in general is not to voice antiblack attitudes or preferences; thus individuals who support the white candidate may be less likely to admit this to either a white or a black interviewer.[10] Thus, testing for race-of-interviewer effects can have important diagnostic implications for assessing the racial tenor of a given contest and for improving predictions of electoral outcomes.

References

Aldrich, John H., and Forrest D. Nelson. 1984. *Linear Probability, Logit, and Probit Models*. Beverly Hills, CA: Sage Publications.

Anderson, Barbara A., Brian D. Silver, and Paul R. Abramson. 1988a. "The Effects of Race of the Interviewer on Measures of Electoral Participation by Blacks in SRC National Election Studies." *Public Opinion Quarterly* 52:53–83.

———. 1988b. "The Effects of the Race of the Interviewer on Race-related Attitudes of Black Respondents in SRC/CPS National Election Studies." *Public Opinion Quarterly* 52:289–324.

Bradburn, Norman, and Seymour Sudman. 1988. *Polls and Surveys*. San Francisco: Jossey-Bass.

Campbell, Bruce. 1981. "Race-of-Interviewer Effects among Southern Adolescents." *Public Opinion Quarterly* 45:231–44.

Clymer, Adam. 1989. "Election Day Shows What the Opinion Polls Can't Do." *New York Times*, November 12.

---

[10] See Sussman (1985) for a discussion of this possibility in the 1985 Virginia election for lieutenant governor and Clymer (1989) for an overview of polling problems in the 1989 election.

Cotter, Patrick, Jeffrey Cohen, and Philip B. Coulter. 1982. "Race-of-Interviewer Effects in Telephone Interviews." *Public Opinion Quarterly* 46:278–84.

Crespi, Irving. 1988. *Pre-election Polling: Sources of Accuracy and Error*. New York: Russell Sage Foundation.

Fowler, Floyd. 1984. *Survey Research Methods*. Newbury Park, CA: Sage Publications.

Guterbock, Thomas M., and Steven E. Finkel. 1989. "A Bad Day for the Pollsters." *Richmond Times-Dispatch*, November 12.

Hatchett, Shirley, and Howard Schuman. 1975–76. "White Respondents and Race-of-Interviewer Effects." *Public Opinion Quarterly* 39:523–28.

Hugick, Larry. 1990. "Polls during the Past Decade in Biracial Election Contests." Paper presented at the annual meeting of the American Association for Public Opinion Research, Lancaster, PA.

Keeter, Scott. 1990. "Race-of-Interviewer Effects in the 1989 Virginia Gubernatorial and State Legislative Election Polls." Paper presented at the annual meeting of the American Association for Public Opinion Research, Lancaster, PA.

Meislin, Richard. 1987. "Racial Divisions Seen in Poll on Howard Beach Attack." *New York Times*, January 8.

Morin, Richard. 1989. "Polling—in Black and White." *Washington Post*, November 5.

Roll, Charles W., and Albert H. Cantril. 1980. *Polls: Their Use and Misuse in Politics*. Cabin John, MD: Seven Locks Press.

Sabato, Larry J. 1991. *Virginia Votes, 1987–1990*. Charlottesville: University of Virginia Center for Public Service.

Schuman, Howard, and Jean M. Converse. 1971. "The Effect of Black and White Interviewers on Black Responses." *Public Opinion Quarterly* 35:44–68.

Schuman, Howard, and Shirley Hatchett. 1974. *Black Racial Attitudes: Trends and Complexities*. Ann Arbor, MI: Institute for Social Research.

Schuman, Howard, Charlotte Steeh, and Lawrence Bobo. 1985. *Racial Attitudes in America: Trends and Interpretations*. Cambridge, MA: Harvard University Press.

Schuman, Howard, and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys*. New York: Academic Press.

Shapiro, Walter. 1989. "Breakthrough in Virginia." *Time*, November 20.

Silver, Brian D., Barbara A. Anderson, and Paul R. Abramson. 1986. "Who Overreports Voting?" *American Political Science Review* 80:613–24.

Sudman, Seymour, and Norman M. Bradburn. 1974. *Response Effects in Surveys*. Chicago: Aldine.

Sussman, Barry. 1985. "Hidden Racial Attitudes Distorted Va. Polls." *Washington Post*, November 28.

Troldahl, Verling C., and Roy E. Carter, Jr. 1964. "Random Selection of Respondents within Households in Phone Surveys." *Journal of Marketing Research* 1:72.

WORK AND LEISURE
ON THE REPORTING OF POLL RESULTS

RICHARD F. HAMILTON

Abstract Louis Harris has reported a substantial increase in work between 1973 and 1985 and a corresponding decrease in the availability of free time. Data from the Current Employment Statistics of the Department of Labor show a decline of the work week in the same period. Data from the National Opinion Research Center's General Social Surveys show a rough constancy in the work week. Previous analyses of Current Population Survey results also point to constancy as the appropriate conclusion for the period in question. Examination of the Harris studies shows five changes in procedure. These changes in method may well have produced the dramatic changes he reported. Harris's finding is challenged also by other results from his studies, which show increased participation in a wide range of leisure-time activities.

Louis Harris, the noted public opinion expert, has presented two remarkable findings in his recent book, *Inside America* (1987). They are introduced in a chapter entitled "Is Leisure Time on the Way Out in America?" The findings are stated clearly, unambiguously, and with no qualification: "Since 1973 the number of hours worked by Americans has increased by 20%, while the amount of leisure time available to the average person has dropped by 32%" (p. 17). Adding some details, a second summary statement has it that "the number of hours the average American spends at work each week has increased from 40.6 in 1973 to a current 48.8, a rise of 8.2 hours a week or one hour and 10 minutes a day" (p. 19). The end point for the period under discussion was 1985; the 20 percent increase in work, therefore, occurred within the relatively brief span of 12 years.

Harris is certainly justified in referring to these as "dramatic shifts in working hours and leisure time." That change in the work week would be the most important development since the setting of the 40-hour standard with the Fair Labor Standards Act of 1938. The re-

RICHARD F. HAMILTON is professor of sociology and political science at Ohio State University.

Public Opinion Quarterly Volume 55:347–356 © 1991 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/91/5503-0008/$2.50

cent change, so it would appear, has undone that prior achievement. One might easily anticipate an increase in reports of stress and dissatisfaction (with work, with free time, with the quality of life), and, accordingly, one of the major themes of *Inside America* is the high level of stress.

One might expect considerable public discussion, and much complaint, about this negative trend, one that has reversed the long-term historical shift toward greater leisure. But the attention paid the findings, by and large, has been limited. An article in the *New York Times* (Yarrow 1988) reviewed Harris's previous results and added figures for 1987 that show the work week still at a high 46.8 hours. The decline in leisure continued, being a mere 16.6 hours in 1987, the lowest point since the original 1973 study. A 1989 account in *Psychology Today* (Harris and Trotter 1989) reports Harris's figures but, curiously, does not indicate the source. The article recounts the finding and also gives a reading of the public reaction. "Work has become our intoxicant," the authors write, "and Americans are working harder than ever before" (Harris and Trotter 1989, p. 33). They speak of "this new passion for work" and quote Daniel Yankelovich, a "leading researcher on changing American values and corporate editor of *Psychology Today*," who declared that "the job is now the center of excitement in American lives."[1] Working with Harris's figures, these authors report that "the typical adult's leisure time has shrunk by 40%—down from 26.6 to 16.6 hours a week."

Although of obvious importance, these dramatic changes have, as indicated, received little attention either in the popular press or in the scholarly literature. One academic specialist, John D. Owen, makes no reference to the Harris findings in two recent publications (Owen 1988, 1989).

It is useful, before considering causes and implications, to take a closer look at "the facts." From easily available sources, we have three options. They are: first, Harris's data showing the dramatic increase in the work week; second, the Department of Labor's Current

---

[1] A minor difference appears in one of the figures given. All others, however, are identical with those contained in the Harris (1987) and Yarrow (1988) accounts. George Harris and Robert Trotter (1989) open their account with a reference to one of the major "hypes" of the modern age, the expectation in the 1950s that automation was going to give rise to "landslide unemployment," a need for job-sharing, and vast quantities of leisure. That claim was ultimately recognized as a mistake and abandoned. They pass over another major "hype," the alarm sounded in the early seventies about a decline in work motivation that, supposedly, came with the new values of self-fulfiliment. Anger and outrage were predicted; there was going to be a "rebellion in the workplace." For a review and dissection of that episode, see Hamilton and Wright (1986, pp. 19–33, 65–95). Yankelovich was one of the many authors arguing those claims (quoted in Hamilton and Wright 1986, pp. 10–11). It is something of a surprise, therefore, to find him now declaring "the job" as "the center of excitement."

Table 1. Numbers of Hours Worked per Week

	Research		
Year	Harris	CES	GSS
1973	40.6	36.9	39.9
1975	43.1	36.1	39.0
1980	46.9	35.3	41.0
1984	47.3	35.2	40.5
1985	48.8	34.9	41.2

SOURCES.—Harris (1987, pp. 17–19); for the CES, see U.S. Department of Labor (1989, p. 79); and for GSS, see Davis and Smith (1987); my calculations.

NOTE.—The Harris figures are medians; the other two series are means. See n. 7 below for further comment.

Employment Statistics (CES) series for the same period showing a modest decrease in the work week (U.S. Department of Labor 1989); and third, results from the National Opinion Research Center's General Social Surveys (Davis and Smith 1987), which show undramatic stability. (See table 1.) Some simple questions arise: Which series is accurate? Has the work week increased, decreased, or remained the same? How may one account for those differences?

Harris's findings come from national surveys undertaken for the Associated Councils of the Arts (ACA), which have been published in a series entitled *Americans and the Arts*.[2] The first survey dates from January 1973. It involved personal interviews with a "representative cross section of all Americans 16 years and older." The sample was "scientifically drawn" and 3,005 persons were interviewed (ACA 1975, p. 3). The key question, together with an important instruction, reads as follows:

> 4a. About how many hours a week does your job or occupation take, including travel time? (INTERVIEWER NOTE: INCLUDE STUDENTS AND HOUSEWIVES IN THIS RESPONSE.) [ACA 1975, p. 131]

Important methodological changes appear in later Harris studies. The 1975 study report provides only sparse details on methods. The only obvious differences are in the date of interview (June) and in the sample size (N = 1,555) (ACA 1976, p. i). Unfortunately, the 1975 questionnaire was not reproduced nor were the key question wordings provided. For these reasons, the principal comparison here will be between the 1973 and 1980 studies.

---

[2] The name of this organization was later changed to American Council for the Arts. It is indicated as the sponsor of the later 1980 survey.

The procedures used in these studies differ in five respects. First, there was a shift from the personal in-home interviews used in 1973. The "in-depth survey" of 1980 was conducted by telephone. The telephone sampling made use of exchange numbers plus random digit dialing (ACA 1981, p. 55).[3] Second, there was a change in the final step of the sampling procedure. In the 1973 study interviewers were required to list all appropriate persons in the household by age and, through predetermined arrangement, were told which person to interview (ACA 1975, p. 125). The description of "methodology" for the 1980 sample reviews all steps down to the selection of seven-digit telephone numbers (ACA 1981, app. B, p. 55). The final step of the procedure is indicated only on the first page of the questionnaire (ACA 1981, app. C, p. 57). A screening question excluded those under age 18. If the person answering was 18 or over, the interviewer was instructed to "continue with questionnaire." If the person answering did not meet this minimum age requirement, he or she was asked "if anyone" there was 18 or over, and the interviewer was instructed to ask for "that person." The final step, in short, involved nonrandom procedure.

Third, the lower age limit of those surveyed in 1973 was 16. In 1980, as just indicated, the lower limit was 18. Fourth, there was a change in the timing of the interviews. The 1973 study was conducted in January. The 1980 study was conducted in July, specifically "between July 17th and 28th." This is the only change, of the five, that is reported in *Inside America* (Harris 1987, p. 21).

Finally, the question used in the 1980 survey was different. The 1980 question reads:

> 1a. First, we would like to know approximately how many hours a week you spend at your job or occupation, and that includes keeping house or going to school as well as work for pay or profit. How many hours would you estimate you spend at work, housekeeping, or studies, including any travel time to and from the job or school? [ACA 1981, p. 58]

This change in wording, by itself, could easily have generated the "dramatic" result. It certainly invites a more generous reading of "work" than was the case with the original question.

Harris's 1987 report, one may conclude, is seriously misleading. Most people, as a result of long training, are likely to think "labor force" or "paid employment" when coming across discussions of

---

[3] The shift to telephone interviews in 1980 would mean the omission of 7 percent of U.S. households (U.S. Bureau of the Census 1986, p. 531).

"work" and "the work week." Significant variations from that practice should be signaled. If the definition of "workers" and the research procedures had been constant, one could make valid judgments about the trend. But the five changes in procedure effectively vitiate any serious use of these surveys for discussion of shifts over time. It is not immediately obvious what impacts these changes might have had. One conclusion, however, is entirely appropriate: ceteris paribus is not at all likely.[4]

The CES series (U.S. Department of Labor 1989), as noted, shows an opposite trend, a linear decrease in the work week. A glance at earlier figures shows this to be a continuation of a long-term trend. The work week figures given for 1965 and 1970, respectively, were 38.8 and 37.1.

These figures drawn from Current Employment Statistics program are based on employer reports submitted by a national sample of firms (U.S. Department of Labor 1989, pp. 134, 138 ff.). This series does not report on the entire labor force—it covers "production or nonsupervisory workers on private nonagricultural payrolls," thus omitting supervisors, managers, and self-employed persons. One may, however, take the results as providing a reasonably accurate picture of the actual working hours for the large segment specified. The results shown in table 1, it should be added, average some diverse patterns. A few categories, mining and construction, for example, show a rather stable pattern. Modest declines appear in wholesale trade and finance. The largest declines are in retail trade and service establishments. Those may not, however, represent actual declines but rather compositional shifts, that is, increased use of part-time employees (Owen 1988, p. 42).

The Department of Labor also draws on the Current Population Survey (CPS) (Bregger 1984; Devens 1978) for information on labor force involvement. This source, as is well-known, involves monthly surveys of households with a sample size of approximately 60,000. It provides data not picked up in the employer reports. The latter does not adequately deal with the problem posed by "moonlighting." Persons with two (or more) jobs could be reported by two (or more) establishments without indication that single job-holders are involved. The CPS household survey picks up information about multiple jobholding and, appropriately, links those jobs to a given person. An analysis comparing the two surveys finds, not surprisingly, that the household data, without exception, "showed longer workweeks than the payroll data for the same industry" (Devens 1978, p. 8).

---

[4] In late July many people would be on vacation or out of the country. Those at home might have been working harder, doing the work of absent colleagues. Self-selection is another possibility: persons valuing leisure would opt for vacations; workaholics, those with a "passion for work," would remain at home where, among other things, they would have been more readily available for interviews. For data on the distribution of vacations by month, see Hedges and Taylor (1980, p. 9).

At least three detailed analyses of the CPS data have reported a rough constancy in the length of the work week for full-time employed persons. The absence of change was reported by Owen (1976). The diverse results—CPS constancy and CES decline—were reviewed by Hedges and Taylor (1980). A careful analysis by Owen, one that focused on full-time workers, "finds virtually no downward movement in the past four decades" (1988, p. 42). None of the three report an increase in the work week. The last two studies, it will be noted, cover periods wherein Harris discovered that "dramatic" increase.[5]

The General Social Surveys (Davis and Smith 1987) are based on multistage probability samples of the United States' noninstitutional adult population (age 18 and over). The samples contain approximately 2,000 persons. Personal interviews are used. Interviews were conducted in February, March, and April in all studies. The work-week question has been constant throughout. It reads:

> Q. [IF WORKING, FULL OR PART TIME] "How many hours did you work last week, at all jobs?"

The category, effectively, is the civilian labor force. It is, therefore, larger than the Department of Labor CES category, equivalent to the CPS category, but smaller than the inclusive Harris category. The figures are all higher than the Department of Labor CES equivalents. Since these include the managers and self-employed—categories with high average work weeks—this is the appropriate result. The GSS figures include, moreover, time spent on second and third jobs.

The best single conclusion about the trend in the work week since 1973, if one were to accept the focus on the civilian labor force, is the third of the indicated options, that of "no change." That is, to be sure, an overall result, one which doubtlessly "merges" many important variations within that total.

One could test the Harris claim by making predictions about likely consequences and assessing those derived claims. As noted, Harris does signal growing stress as a consequence of this twofold process, more work and less leisure. Harris (1987) does make some exploration of this possibility through use of the National Opinion Research Center's General Social Surveys (pp. 40–41). Comparing the seventies and eighties (the precise dates are not indicated), he reports declines in satisfaction with friendships (from 70 to 65 percent), with health and physical condition (from 61 to 55 percent), and with nonwork, leisure-time activities (from 56 to 52 percent). Those results would certainly be consonant with his initial work and leisure findings. It is surprising, however, that the "dramatic" 32 percent decline in leisure time should yield only a four-point reduction in satisfaction with leisure. All or part of those small changes, moreover, could be due to "other causes." In an aging population, for example, a decline in satisfaction with health and physical condition is a "normal" expectation. A loss of friends is also a normal correlate of advancing age.

---

[5] Current Population Survey data indicate an increase but only a small overall incidence of "moonlighting." Recent data show about 6 percent of employees report having more than one job (Owen 1988, p. 42; Stinson 1986, 1987). Because of compositional changes (increases in women's employment, part-time work, and student involvement), trends are difficult to establish. Owen (1988) has presented CPS data for nonstudent, full-time employed men. That segment was once dominant in the labor force and even now accounts for about half of the total. His data show "virtually no downward movement" in the past 4 decades—and no increase. His figures for the period in question, in weekly hours are: 1972, 42.8; 1975, 42.2; 1978, 42.7; 1981, 42.2; 1984, 42.5, and 1986, 42.8. These figures are all slightly higher than the GSS figures shown in table 1. Employed adult men, Owen reports, work "about 7 hours more a week than do employed women." Both work many more hours than the employed students (Owen 1988, p. 42).

The Harris reader familiar with the General Social Surveys would notice two omissions (Harris 1987, pp. 39–40). Harris has not reported GSS trend figures for the "happiness" question or for the job satisfaction question. Only 29 percent of the sample described themselves as "very happy" in 1985. That was the lowest figure in the entire series which began in 1972. But that is at best a mini-finding since all figures from 1972 through to 1988 vary only slightly, the extremes being 29 percent and 38 percent. The averages of the first 4 and the last 4 years in the series were, respectively, 34 percent and 32 percent. With respect to work satisfaction, the decline is larger, a matter of three percentage points. In short, those measures of stress—the happiness and work satisfaction questions—yield only very modest support for Harris's principal claims.

Still another possibility is to look at internal evidence, that is, at other items contained within Harris's own account. The discussion of the work week, as noted, appears first on page 17, this under the heading "Leisure Time on the Way Out in America?" On page 159, one finds an entirely unexpected heading: "Mass Need for Creative Expression: The Explosion of the Arts." In the following paragraphs Harris reports that "from the mid-1970s to the mid-1980s, arts attendance in this country literally exploded." Attendance at movies increased as did attendance at "live performances of plays, musical comedies, pantomime, and other forms of theater." Similar increases are reported for live music performances by popular singers, bands, and rock groups. Visits to art museums increased marginally. Attendance at opera or musical theater performances increased significantly. The same is reported for ballet, modern dance, and folk or ethnic dance performances. Attendance at performances of symphony and chamber orchestras also increased. Harris reports also that sizable portions of the population "feel a need for more presentations of a host of arts fare." In addition to the passive viewing or listening, Harris devotes a couple of pages to direct involvement—"Since the mid-1970s, participation in the arts has increased even more." All of these findings stem from the *Americans and the Arts* surveys, which yielded the initial, now very problematic, finding on work and leisure.

Harris's findings on work and leisure appear to be a function of method. Question 1a in the 1980 questionnaire was the sweeping question about "work." Question 1b was a parallel "how many hours" question inquiring about total leisure time.[6] It seems likely that respondents would have supplied appropriate answers; if one gave a high figure for "work," that would have been balanced by an appropriate low "leisure" figure. Question 2 listed a wide range of leisure activities and asked "whether you generally do it a great deal, some but not a great deal, very little, or not at all." Harris's respondents may, at all points, have provided accurate responses to all of these questions. Reasonable persons would respond differently to the different questions contained in the 1973 and 1980 studies, thus giving rise to the appearance of increased work and decreased leisure. The 1980 respondents could, at the same time, have provided accurate reports about the frequency of their leisure time activities, which yielded the apparently "contradictory" finding of an "explosion" in leisure.[7]

The following conclusions appear warranted. There was no dramatic increase in "work" between 1973 and 1985 nor was there a dramatic decrease in leisure. The best judgment on the former question is: rough constancy of the workweek (of those in the labor force). This is indicated by the Current Population Surveys (adjusted for compositional changes) and the General Social Surveys. These, on balance, appear to provide the best measures among the options considered here. The Harris workweek "finding" appears to reflect changes in the methods used rather than any real change. The same conclusion appears justified with respect to his finding about leisure time.[8]

---

[6] A minor detail is that the Harris question on "work" reminds the respondent to include "any travel time." No comparable reminder appears in the leisure question.

[7] My letter of inquiry addressed to Louis Harris was answered by Matthew Goldstein, president of the Research Foundation of the City University of New York (October 11, 1989). Among other things, Goldstein noted that the figures given in *Inside America* (Harris 1987) are medians. The 48.8 figure for 1985 (shown here in table 1) is the median number of "work" hours for 1985; the mean, he reports, was 44.9. The original Harris publication gave the 1973 "mean number of working hours per week" as 34.9 (ACA 1975, p. 66). The table there also gives the percentage of respondents in all categories. The largest segment, 23 percent, was the "Do not work" category. Harris's *median* figure (reported here in table 1) appears to exclude that segment. The figure given for 1973 leisure—the "mean number of hours"—was 24.8 (ACA 1975, p. 67). That differs slightly from the figure given in his 1987 publication, 24.3, but that, according to Goldstein, is the median. It is not clear how the "not at work" segment is handled there, whether included or excluded in the calculation of leisure time. In the 1975 publication they were clearly included. The report of the 1975 study (ACA 1976, p. 78) gives the "average hours per week of work and leisure time" as, respectively, 38.8 and 25.3. The medians (Harris 1987, pp. 17, 19) are given as 43.1 and 24.3, respectively.

References

Associated Councils of the Arts. 1975. *Americans and the Arts: A Survey of Public Opinion*. New York: ACA. (Report of the original 1973 survey.)

———. 1976. *Americans and the Arts: A Survey of the Attitudes toward and Participation in the Arts and Culture of the United States Public*. New York: ACA. (Report of the 1975 survey.)

American Council for the Arts. 1981. *Americans and the Arts: A 1980 Survey of Public Opinion*. New York: ACA. (Report of the 1980 survey.)

Bregger, John E. 1984. "The Current Population Survey: A Historical Perspective and BLS' Role." *Monthly Labor Review* 107(6):8–14.

Davis, James Allan, and Tom W. Smith. 1987. *General Social Surveys, 1972–1987* (machine-readable data file). Chicago: National Opinion Research Center.

Devens, Richard M., Jr. 1978. "The Average Workweek: Two Surveys Compared." *Monthly Labor Review* 101(7):3–8.

Hamilton, Richard F., and James D. Wright. 1986. *The State of the Masses*. New York: Aldine.

Harris, Louis. 1987. *Inside America*. New York: Vintage.

Harris, T. George, and Robert J. Trotter. 1989. "Work Smarter, Not Harder." *Psychology Today* 23:33.

Hedges, Janice Neipert, and Daniel E. Taylor. 1980. "Recent Trends in Worktime: Hours Edge Downward." *Monthly Labor Review* 103(3):3–11.

Owen, John D. 1976. "Workweeks and Leisure: An Analysis of Trends, 1948–75." *Monthly Labor Review* 99(8):3–8.

---

[8] A recent publication has reviewed another discrepant result, this involving Harris's 1984 *America and the Arts* study (cited in Robinson [1989]). The Harris study reported considerably higher levels of arts participation than were found in a study conducted by the U.S. Bureau of the Census (cited in Robinson [1989]). The Harris sample overrepresented, by considerable margins, persons with high education and high income. Persons in those categories, Harris noted, report longer than average "work" hours and less than average leisure (1987, pp. 19–20). Yarrow (1988) also reported discrepancies between Harris's findings and those of other researchers. The 1987 Harris study found the lowest amount of leisure time yet recorded and, at that point, also found a decline in arts attendance. Studies of concert attendance, museum going, and motion picture admissions all reported increases. At first sight, the increase in leisure-time activities seems difficult to reconcile with the constancy of the workweek indicated by the General Social Surveys (Davis and Smith 1987). But life consists of more than work and leisure. It might be that mealtimes have suffered. Fast foods and diet menus may have "freed up" time for those other things. The Yarrow article (1988), it will be remembered, reports Harris's recent finding—only 16.6 hours of leisure per week for the average adult American. That would mean about 2.4 hours per day (or possibly 2 hours on weekdays with 3 and 4 on weekends. Were that the case, it would give rise to another perplexing question. How would the "average American," who is so heavily burdened with "work," be able to fit in that often-reported 6 hours per day of television viewing? That too is a fiction. To achieve that result, the average American would have to watch television, e.g., from 6 P.M. to midnight. The General Social Surveys routinely report a median of 2 hours.

———. 1988. "Work-Time Reduction in the U.S. and Western Europe." *Monthly Labor Review* 111(12):41–45.

———. 1989. *Reduced Working Hours: Cure for Unemployment or Economic Burden?* Baltimore: Johns Hopkins University Press.

Robinson, John P. 1989. "Survey Organization Differences in Estimating Public Participation in the Arts." *Public Opinion Quarterly* 53:397–414.

Stinson, John F., Jr. 1986. "Moonlighting by Women Jumped to Record Highs." *Monthly Labor Review* 109(11):22–25.

———. 1987. "Moonlighting: A Key to Differences in Measuring Employment Growth." *Monthly Labor Review* 110(2):30–31.

U.S. Bureau of the Census. 1986. *Statistical Abstract of the United States: 1987*. 107th ed. Washington, DC: Government Printing Office.

U.S. Department of Labor, Bureau of Labor Statistics. 1989. *Employment and Earnings* 36(9):79, table C-1.

Yarrow, Andrew L. 1988. "Harris Poll Finds Arts Attendance Has Declined." *New York Times*, March 16.

VOTERS' INTERMEDIATION ENVIRONMENTS IN THE 1988 PRESIDENTIAL CONTEST

PAUL ALLEN BECK

Abstract Considerable information about a presidential election campaign comes to voters through such intermediaries as personal networks and the mass media. This article examines the operation of these intermediaries in the 1988 Ohio presidential campaign by focusing on exposure to them and the perceived partisan content of their messages—and how these key features of intermediation relate to various voter characteristics. People are most disposed to discuss politics with like-minded relatives and friends, but discussions with co-workers provide an opportunity for politically dissonant messages to intrude. This political discussion is modulated by political attentiveness rather than personal or partisan characteristics. Media exposure is high and also is related most to political attentiveness, although age is an important determinant as well. Surprisingly, the media were commonly perceived as balanced and neutral in the presidential race or, where committed, as supporting the candidate the respondent opposed. Because of this, only a third of the respondents found themselves in an overall intermediation environment that was concordant with their own preferences. Considerably more were in environments containing some discordant views, which could challenge their partisan dispositions.

Citizens in modern democracies have little direct contact with politics, especially at the national level. For most, the information from which cognitive maps of politics are constructed and on which political decisions are based comes almost entirely through intermediaries between them and an often remote political world (Nimmo and Combs 1983).

PAUL ALLEN BECK is professor of political science at Ohio State University. This is a revised version of a paper presented at the 1989 annual meeting of the American Political Science Association in Atlanta on September 3, 1989. The author is grateful to the Polimetrics Laboratory at Ohio State University for conducting the survey and to Russell Dalton, Robert Huckfeldt, John Kessel, Jon Krosnick, Bradley Richardson, Steve Weatherford, and two *Public Opinion Quarterly* referees for their constructive suggestions.

Public Opinion Quarterly Volume 55:371–394 © 1991 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/91/5503-0003/$2.50

Knowledge about voter exposure to various intermediaries and about the nature of the political messages the intermediaries convey is essential for understanding citizen political behavior in modern societies.

Contemporary political life is dominated by two types of intermediaries—personal networks and the mass media. Personal networks include the people who are in face-to-face contact with the individual—family, friends, neighbors, and co-workers. For most Americans, these personal intermediaries are a critical source of information about politics and of political evaluations, as they probably always have been. The mass media are a newer source of political information. They have become increasingly important in this century with the growth of the "popular press" and the penetration of television. By the 1960s, many observers concluded that television had emerged as the dominant political medium in the United States, joining newspapers and displacing radio as the primary sources for political information.[1]

In view of the acknowledged importance of these intermediaries, it is surprising that over the years so little research has focused on their role in individual political behavior. The seminal research on the topic by Paul Lazarsfeld and his colleagues at Columbia University (Berelson, Lazarsfeld, and McPhee 1954; Katz and Lazarsfeld 1955; Lazarsfeld, Berelson, and Gaudet 1948) is drawn from data now 4 decades old that predate the emergence of television, the social and geographical mobility of the postwar period, the increasing political awareness of women, and the recent decline of partisanship—to name but a few recent changes that may have transformed the intermediation process.

Since the Columbia studies, research on political behavior has been rooted in intellectually competing social structural and individualistic approaches, neither of which focuses directly on the social process of intermediation. The social structural approach assumes individual political behavior is mediated by secondary groups but is content to leave this assumption unexamined by largely ignoring the mechanisms of social group influence (Sartori 1969). The individualistic approach views individual political behavior as either psychologically determined (Campbell et al. 1960) or the product of rational self-interest calculations (Fiorina 1981), but these contrasting theories share a conception of individuals as atomistic political actors, divorced from their social contexts.

Recent years have seen a resurgence of research on personal and media intermediaries (Eulau 1980; Sheingold 1973). The role of personal intermediaries has been illuminated by extensive empirical explorations of personal networks (Huckfeldt and Sprague 1987; Laumann 1973). The effects of the mass media on citizen political behavior have been examined systematically through experimental research on communication effects and survey studies of media audiences (Kinder and Sears 1985).

---

[1] A third type of intermediary, secondary organizations, was measured in the survey but will not be considered here.

For all of their value in moving away from treatment of the individual as an autonomous political actor and in estimating the effects of particular intermediaries, however, this recent research runs in separate streams—concentrating on personal networks or television or newspapers to the exclusion of the other intermediaries. The individual's intermediation context, however, is the joint product of all of their intermediaries, each exerting its own unique kind of influence and perhaps, as Chaffee (1982) has suggested, complementing the other. To understand how the process of intermediation shapes political outlooks and behavior, it is imperative that personal and media intermediaries be taken into account simultaneously.

This study is designed as a first step in a comprehensive examination of the role of intermediaries by studying the intermediation process through the eyes of the individual as it involves personal networks and the mass media. The data were collected through a telephone survey of the Ohio electorate immediately after the 1988 presidential election.[2] The interviews probed perceptions of both exposure to and the partisan favoritism of personal and mass media intermediaries during the presidential campaign. Exposure and favoritism are key dimensions of intermediation, and any comprehensive study of intermediation must focus on them.

---

[2] Interviews were conducted from November 9 through December 15, 1988, from Columbus, Ohio, by paid, trained, and professionally supervised interviewers. The study sample was drawn by (1) selecting a random sample of numbers from the state's telephone directories; (2) substituting random digits for the last two digits of the chosen number (Frey 1989; Sudman 1983); and (3) randomly selecting an adult respondent at that number. Of 1,288 eligible numbers dialed, 652 were interviews, 311 were refusals, 217 went unanswered after more than six calls at different times and on different days, and 108 were unanswered with under six callbacks. This produces an unadjusted response rate of 652/1,288 = 51 percent. But this number underestimates the real response rate, because many unanswered numbers are actually nonworking numbers. Two approaches seem reasonable to adjust for this. If all callbacks after six tries are treated as ineligible, the denominator of the response rate is reduced by 217, increasing the response rate to 652/1,071 = 61 percent. This compares favorably to National Election Study (NES) telephone rates in the 1984 rolling cross-section (62 percent) and 1988 Super Tuesday studies (63 percent) in which all nonanswers called 14 or more times were eliminated. (Only a handful of additional NES contacts were made after six calls.) A more conservative calculation would reduce the denominator by excluding only the likely ineligibles from the nonanswers. Of all numbers for which a disposition was determined, ineligibles were 54 percent. This works out to 100 nonanswered numbers as eligibles, or a response rate of 652/1,171 = 56 percent. The actual response rate for the survey probably falls somewhere between 56 and 61 percent.

Personal Networks in the Intermediation Process

Even though it is cast in the privacy of the voting booth, the vote may be seen as the result of social activity through which people develop, activate, and reinforce their preferences in discussions with others during the course of an election campaign. How much do people discuss politics, and what motivates their discussion practices? How much does this discussion expose people to different political viewpoints? Our data offer some suggestive answers to these important questions.

Exposure. In order to answer these questions about personal intermediation, a series of questions was posed to each survey respondent (ego) about their discussions of the 1988 campaign with others (alter). (Question wording is reported in the Appendix.) About a quarter of the sample (23 percent) reported that they had not talked about the 1988 presidential campaign with anyone, and no further attention is paid to them. The remaining three-quarters were asked specifically about the frequency of their political discussions with spouses, neighbors, and co-workers and with the one person with whom they discussed politics the most.

Respondents were most likely to discuss the 1988 campaign with their spouses and co-workers and least likely to have conversations with neighbors (17 percent). Among those with the opportunity for political interaction with spouse (50 percent) and co-workers (54 percent), political discussion was widespread. In all, 62 percent of married respondents discussed politics with their spouses, and 55 percent of those working outside the home discussed politics with fellow workers.

More insight into these patterns of discussion can be gained by identifying the alter talked with most about the election. Most mentioned someone within their family as their main discussant: 42 percent cited spouses and another 26 percent other relatives, such as parents and adult offspring. Friends and co-workers were named as main discussants by 15 percent and 14 percent, respectively, while only 1 percent named their neighbors.

Variations in personal exposure. How exposure to the various intermediaries varies by individual characteristics can tell us a great deal about the intermediation process. To explore this matter, the relationships of exposure to 10 different personal characteristics were calculated. Four are standard demographic variables—race, gender, age, and education—that locate the individual in the social structure. The remaining variables reflect political attitudes. Political attentiveness is measured by interest in the 1988 campaign, concern about who would win the 1988 presidential election, and how well-informed about politics the interviewer felt the respondent was.[3] Partisan commitment is represented by the strength of liberal or conservative ideology and Democratic or Republican partisanship and the perceived distance between the candidates.[4]

The relationships between these 10 exogenous variables and exposure to personal intermediaries are presented in table 1. The bivariate correlations are presented to permit comparisons with previous studies, but our inferences about intermediation will be based on the regression coefficients.

The political attentiveness variables are the most important correlates of exposure to personal intermediaries. These relationships suggest a motivational basis to exposure, although the possibility of a reverse causal direction cannot be dismissed entirely. That attentiveness is tied to discussion has an important implication for individual political change. The least attentive citizens typically interact less with personal intermediaries. This isolates them from the stream of political discourse accompanying a campaign and undoubtedly makes them more vulnerable to each piece of information about the campaign (Converse 1966).

The political attentiveness variables bear the strongest relationships to spouse and main discussant exposure, the alters consciously chosen by the respondent. No other variables in the model are significantly related to discussion with the spouse. Where the main discussant is concerned, however, two demographic characteristics—gender and age—join attentiveness. Younger adults and females are significantly more exposed to a main discussant.

Different forces appear to structure campaign discussions with neighbors and co-workers. Attentiveness, operating via information levels, is tied to discussion with co-workers, but it loses its primacy to age. The young are more inclined than their elders to enter into political discussions at work—a phenomenon that will receive detailed treatment later. Also, men are somewhat more likely than women to talk politics in the workplace, partly because more men are employed outside of the home.

The relationships between age and discussant exposure are more complex. There is a precipitous downturn in political discussion with nonrelatives after age 60, as the physical isolation that often accompanies retirement and aging restricts opportunities. By contrast, adults

---

[3] Interviewer assessments have been shown to be the best single measure of how well-informed respondents are about politics (Zaller 1985).

[4] The strength measures are derived from folding their self-identification scales at the scale's midpoint. Candidate commitment is measured by the absolute difference between thermometer evaluations of the two candidates.

Table 1. Characteristics of Exposure to Personal Intermediaries

	Spouse		Neighbor		Co-worker		Main Discussant	
	r	b	*r*	b	*r*	b	*r*	b
Demographic characteristics:								
Race	−.07	−.14	.04	.08	.04	.09	−.06	−.14
Gender	−.03	−.09	−.06	−.06	.12	.12	−.09	−.17
Age	.04	.00	.02	.00	−.28	−.01	−.12	−.01
Education	.13	.02	.09	.02	.23	.03	.18	.03
Political attentiveness:								
Campaign interest	.20	.08	.15	.05	.18	.08	.29	.17
Concern about outcome	.21	.13	.09	.00	.15	.05	.24	.07
Information	.28	.17	.15	.02	.27	.16	.31	.17
Partisan commitment:								
Ideological strength	.13	.06	.01	−.01	.09	.00	.11	.02
Party strength	.04	−.04	.11	.03	.07	.01	.14	.03
Absolute candidate distance	.08	.00	.07	.00	.07	.00	.20	.00
Adjusted R^2		.10		.03		.16		.17
Cases	560		561		560		559	

NOTE.—The entries in each column are Pearson correlations (*r*) and multivariate unstandardized regression coefficients (b) for each row variable and exposure (scored 0 = none to 3 = many times). Higher scores on personal characteristics are assigned to blacks, males, elderly, college educated, interested in the campaign, concerned about the outcome, well-informed, strongly ideological, strongly partisan, and candidate thermometer ratings that are farthest apart. Coefficients significant at the .05 level are underlined.

under age 30 are more likely to have political discussions outside of the home, especially on the job, and to choose nonrelatives as their main discussants. The young-adult life-style fosters interaction beyond the family, especially among unmarried young people. This may increase their exposure to dissonant political cues and, consequently, to challenges to their own preferences—a hypothesis warranting further testing.

One has to go beyond the relationships reported in table 1 to a differentiation in choice of main discussants in order to find the gender differences expected on the basis of previous research. Not only did men have significantly less exposure to a main political discussant but they were also less likely to draw their main alters from the family. Among those identifying a main discussant, about 76 percent of the women chose relatives in comparison with only 56 percent of the men—a significant difference of 20 percent ($r = -.20$). This tendency for women to focus their political discussions within the family seems to have withstood considerable change in the role of women and their exposure to the world outside of the family, as it was found in the 1940s (Berelson, Lazarsfeld, and McPhee 1954) as well as the 1980s (see also Brickell, Huckfeldt, and Sprague 1988; Huckfeldt and Sprague 1987). Their differential in selection of main discussants, then, exposes men more than women to forces for political change.

Concordance between ego and alter. Exposure subjects people to the viewpoints of others, but it is the content of these viewpoints that challenges or reinforces orientations. During a presidential campaign, the candidate preferences of alters is presumably of greatest import, so perceptions of it are used as the measure of alters. Table 2 shows the extent of agreement in presidential preferences between ego and alter for each type of intermediary.[5]

The most striking result in table 2, although it is hardly unexpected (Berelson, Lazarsfeld, and McPhee 1954; Huckfeldt and Sprague 1987; Laumann 1973), is how rarely respondents see themselves in a discordant personal network. For spouses and main discussants especially, the overwhelming tendency is for the alter to share the respondent's presidential preference. High levels of spousal agreement are doubtless the work of initial similarities in viewpoints so common among married couples plus the pressures toward agreement induced by intimate daily interaction. The high agreement levels with main discussants probably reflect conscious choice of discussion partners along political lines.

More intensive examination of political agreement between respon-

---

[5] Alters not perceived as favoring either Dukakis or Bush were coded as "neutral." For neighbors and co-workers, where discussions may involve one or more individuals, the category assigned was based upon which candidate the alters favored on balance—or if they favored neither one.

Table 2. Concordance between Respondents and Their Personal Intermediaries

	Spouse	Neighbor	Co-worker	Main Discussant
Alter's preference is:				
Concordant	79%	43%	48%	76%
Neutral	13%	41%	36%	7%
Discordant	8%	16%	16%	17%
Ego-alter correlations	.74	.36	.37	.61
Cases	199	76	176	259

NOTE.—This table contains voters and those exposed to each alter only. Cell entries are the percentages who perceive the vote preferences of the column alters to be like theirs, neither pro-Bush nor pro-Dukakis, or different from theirs, respectively. The correlations are τ_c's computed between a dichotomous measure of respondent votes and respondent perceptions that alters are pro-Bush, balanced or neutral, or pro-Dukakis, with the Dukakis alternative given the highest score on each measure. Coefficients significant at the.05 level are underlined.

dents and main discussants reveals an important divide between main discussants who are within and outside the family. Where the main discussant was a family member, either a spouse or some other relative, concordance in presidential preferences in 1988 reached 80 percent (τ_c = .71). For main discussants who were drawn from outside the family, on the other hand, concordance fell off to 69 percent (τ_c = .37). Even though agreement is high in both groups, the tendency toward greater political homogeneity within the family is significant for the intermediation process, especially considering differential tendencies by age and gender to find main discussants outside of the homogeneous family circle.

It perhaps should not be surprising that these tendencies toward political homophyly (i.e., likes associating with likes) do not extend in equal measure to neighbors and co-workers, in spite of the exaggerated similarities inherent in self-reports (Niemi 1974). In less than half of either the neighbor or the co-worker interactions can consonance in presidential preferences be found. It is almost as likely that these environments are neutral—or, given our coding for multiple alters, contain several alters who are split in their presidential choice.

Overall these figures suggest that personal networks of political discussants provide protective cocoons for an individual's political preferences, although the tendencies toward political homophyly perhaps are not as pronounced as previous research suggests. This cocoon is more likely to be penetrated, however, as one's network expands beyond the walls of home and family into the broader world. Political interaction outside the family is more likely than interaction with relatives to introduce discordant political viewpoints to the individual, thereby challenging his or her own political preferences. This illustrates "the strength of weak ties" as identified by Granoveter (1973).

This conclusion, in turn, must be qualified by the complete absence of political discussion with alters for many people. Perhaps failure to discuss politics signifies a conscious withdrawal from potential political disagreement, although given the lack of attentiveness and commitment of nondiscussants, lack of interest in things political seems more likely. Whatever the reason, about a quarter of the respondents were embedded in personal networks they perceived to be silent regarding preferences for the candidates in the 1988 presidential election.

Variations in concordance. Like exposure, political agreement between ego and alter is surely not distributed evenly across the population. The degree to which individuals are concordant or discordant with their personal networks should vary with demographic characteristics, political attentiveness, and partisan commitment. This is investigated in table 3.

Except for main discussants, the demographic, attentiveness, and commitment variables do a better job of differentiating ego-alter agreement than they did for levels of exposure. As before, leading the way in explanatory power is political attentiveness, particularly concern about the outcome and political information. These relationships are negative, though, indicating that attentive respondents are more exposed to discord in their personal networks. Attentives may be more likely to talk about politics wherever they are. Circumstantial evidence for this inference may be drawn from the null relationships: when it comes to their main discussants, attentives are no longer more likely to report dissonance in presidential preferences. All in all, though, the fact that the people most interested in the campaign and best informed about politics—and thereby presumably most capable of resisting the influence of dissonant political messages—are the most exposed to dissonance may limit the influence of heterogeneity in personal networks. Weak ties may not be so strong after all.

Beyond the attentiveness measures, only race and age show a significant relationship in the multivariate analysis—and then only for discussion with co-workers. Perhaps the very one-sidedness of their strongly pro-Dukakis preferences gave blacks less opportunity than whites to encounter consonant presidential preferences among workplace discussants, especially in more integrated settings. Young adults may be more exposed to dissonant preferences in the workplace, on the other hand, because their disproportionate location in unfamiliar

Table 3. Characteristics of Concordance with Personal Intermediaries

	Spouse		Neighbor		Co-worker		Main Discussant	
	r	b	*r*	b	*r*	b	*r*	b
Demographic characteristics:								
Race	−.02	−.87	−.10	−1.45	−.13	−.04	−.05	−.44
Gender	.02	.37	.16	.72	.00	.22	−.07	−.27
Age	−.07	−.02	−.12	−.01	−.22	−.04	.08	.01
Education	−.14	−.08	−.01	−.03	−.18	−.10	−.10	−.10
Political attentiveness:								
Campaign interest	−.17	−.08	−.11	.27	−.30	−.54	−.01	.07
Concern about outcome	−.27	−.72	−.22	−.62	−.33	−.82	−.04	−.02
Information	−.28	−.59	−.24	−.69	−.31	−.56	−.05	−.01
Partisan commitment:								
Ideological strength	−.03	.12	.09	.30	−.09	.00	−.08	−.15
Party strength	−.10	.06	.01	.32	−.07	.22	.00	.04
Absolute candidate distance	−.20	−.01	−.10	−.01	−.21	−.01	−.04	.00
Adjusted R^2		.11		.08		.21		.00
Cases	250		96		239		360	

Note.—This table contains voters and those exposed to each alter only. The column entries are Pearson correlations (*r*) and multivariate unstandardized regression coefficients (b) for each row variable and perceived alter preferences (scored −1 = opposite, 0 = neutral, and 1 = same). Higher scores on personal characteristics are assigned to blacks, males, elderly, college educated, interested in the campaign, concerned about the outcome, well-informed, strongly ideological, strongly partisan, and candidate thermometer ratings that are farthest apart. Coefficients significant at the .05 level are underlined.

job environments in which patterns of interaction have not yet been routinized.

Greater insight into this aspect of the personal intermediation process again can be gained from separating relatives from nonrelatives among the main discussants. Only when this is done do gender differences appear. Men not only were decidedly less inclined to choose relatives as main discussants but also the relatives they chose agreed less with them than was the case for women (59 percent for men versus 75 percent for women).

Important differences in choice of main discussant also appeared between adults under 30 and their elders. Almost twice as many young as older voters (44 percent to 23 percent) chose main discussants from outside of the family. Because extrafamilial main discussants are more likely to hold dissonant presidential preferences (by a 30 percent to 11 percent margin), the greater exposure of the young to them already places young voters in less politically homogeneous political environments. This exposure to different partisan views is heightened by the fact that the nonrelative main discussants of young adults are much more likely to hold dissonant preferences than are the nonrelative main discussants of adults age 30 and older (by a 43 percent to 23 percent margin).

The Mass Media in the Intermediation Process

While personal networks are universally regarded as important political intermediaries, it is the mass media, especially television, that increasingly are seen as the principal actors in the intermediation process (Ranney 1983). Because the media serve as prime conduits for the flow of information from candidates and campaign to voters, they may even be a primary source of information that is exchanged through personal networks. How attentive are people to the media, and who are the attentives? To what degree do the media expose people to contrary political views? These are questions that any study of the flow of information through the mass media must attempt to answer.

Exposure. To explore the use of the mass media for news about the presidential campaign, respondents were asked how much attention they paid to the national evening television news and to daily newspapers, the two principal media for political information. As expected, attention to these media was widespread. Over 90 percent said they watched TV news or read a daily newspaper at least once in a while. Many fewer used these media on a regular basis, though, and, surprisingly, regular usage of newspapers (48 percent) for political news surpassed that of TV news (34 percent).

Variations in media exposure. Just like exposure to personal intermediaries, exposure to mass media for political information may vary by demographic characteristics, political attentiveness, and partisan commitment. Table 4 reports these relationships.

While the 10 variables account for television usage much better than newspaper usage, the characteristics of audiences for the two media are quite similar. Political attentiveness is related to greater usage of both, although a motivational variable (campaign interest) leads the way for television whereas a cognitive measure (information about politics) is more important for newspapers. That these media attract a politically interested and informed audience seems natural given the motivation required to attend to a newspaper or tune in the evening news.

Exposure to both also increases significantly with age. This has important implications for individual political change. Older adults are deeply embedded in familial personal networks, which typically reinforce rather than challenge their political preferences. The protective political cocoon of older adults, however, may be penetrated by their greater reliance on television and newspapers. A total of 67 percent of those age 60 or older reported watching TV news every day compared to only 17 percent in the under-30 group. The comparable figures for reading a newspaper every day are 73 percent for the 60+ group and 28 percent for young adults. What older adults may lack in personal discussions of politics seems to be compensated for by a greater attentiveness to the media. The media's messages may provide the challenges to political preferences that personal networks do not.

Only in the racial characteristics of the two audiences do television and newspapers really diverge. Blacks are significantly more likely than whites to tune in television news programming, but no racial differences emerge in newspaper usage. Faint differences, familiar to media scholars, also appear for education, but they fail to achieve significance in the multivariate analysis. Television news disproportionately attracts people with low levels of education—just the opposite from newspapers.

Concordance between ego and media. The development of the mass media, and their effective propaganda use by totalitarian political regimes in the 1930s and 1940s, initially raised the specter of the media as a potent instrument of political intermediation. Because early research on the media was unable to detect significant effects, however, the all-powerful model of media effects was quickly replaced by its antithesis—what came to be known as the "minimal effects model" (Kinder and Sears 1985; Klapper 1960). This model was supported in the Berelson, Lazarsfeld, and McPhee (1954) study of the 1948 presi-

Table 4. Characteristics of Exposure to Mass-Media Intermediaries

	TV News		Newspapers	
	r	b	*r*	b
Demographic characteristics:				
Race	.09	.38	−.01	.03
Gender	−.10	−.12	.02	.04
Age	.36	.02	.26	.02
Education	−.14	−.03	.04	.03
Political attentiveness:				
Campaign interest	.16	.19	.08	.07
Concern about outcome	.02	−.08	−.01	−.07
Information	.12	.09	.18	.17
Partisan commitment:				
Ideological strength	−.03	−.03	.01	−.01
Party strength	.11	.05	.06	.03
Absolute candidate distance	.06	.00	−.04	.00
Adjusted R^2		.18		.09
Cases	560		561	

Note.—The entries in each column are Pearson correlations (*r*) and multivariate unstandardized regression coefficients (b) for each row variable and exposure (scored 0 = none to 3 = watch news/read newspaper every day). Higher scores on personal characteristics are assigned to blacks, males, elderly, college educated, interested in the campaign, concerned about the outcome, well-informed, strongly ideological, strongly partisan, and candidate thermometer ratings that are farthest apart. Coefficients significant at the .05 level are underlined.

dential campaign in Elmira, New York, and in some of the more recent studies of television (Patterson 1980; Patterson and McClure 1976).

Recent research on news conveyed by the media justifies a refined model of effects that fall between the previously dominant all-or-nothing extremes (Roberts and Maccoby 1985). The media have been found to play an important role in agenda setting and, through it, in priming candidate preferences (Iyengar and Kinder 1987) and framing causal attributions (Iyengar 1987). Recent research also emphasizes how conditional media influence is on characteristics of the audience and the circumstances of the communication (Erbring, Goldenberg, and Miller 1980).

What often goes unrecognized in assessing media effects is how much the American media attempt to be balanced and objective in news coverage of presidential campaigns. For many years, federal regulation curtailed favoritism toward particular candidates on television and radio. Even with the demise of the "fairness doctrine," the electronic media have adhered to their traditional stance of neutrality in news coverage. Although newspapers have been under no such governmental constraints, and originated as instruments for partisan propaganda, they too usually eschew blatant favoritism in news reporting and relegate their support of candidates to the editorial page. The once partisan press has succumbed to the professionalism of modern-day journalism.

While careful partisan observers sometimes claim that they can detect partisan "slants" to the news, systematic empirical research finds objectivity and balance rather than favoritism in both newspapers (Graber 1980) and television (Robinson and Sheehan 1983). Even if there is a partisan slant to the news, it must be so subtle that it fails to register with most readers—and scholars. The result may be a more circumscribed role for the mass media as a political intermediary in presidential contests in spite of its widespread use for gaining political information. The media may be a conduit for information, but they typically provide little partisan slant to the messages they convey.

Evidence for the hypothesis of media neutrality can be found in perceptions of media favoritism in our data. Respondents were asked whether each of the media they used favored a candidate or was balanced between them. As table 5 shows, at least 60 percent of those exposed to each of the media perceived them to be balanced between Dukakis and Bush. The percentage for newspapers is surprisingly high considering that most endorsed a presidential candidate on their editorial page. Also surprising, given the different histories of the two media, are the perceptions that television showed almost as much favoritism as newspapers and was somewhat more likely to favor the opposition candidate.

Among those who perceived media favoritism, moreover, the dominant tendency was for people to perceive television and newspapers as less favorable to their preferred candidate than to his opponent. Can it be that the media consistently support candidates that their viewers or readers oppose? Although possible, that is unlikely. Instead, what we feel we have captured is a psychological tendency to perceive actual partisan balance as favoritism for the candidate one opposes—a posture far easier to adopt for an impersonal institutional source than a personal discussant. This phenomenon is labeled "negative projection" to differentiate it from the more familiar positive projection of personal preferences upon a trusted source. Vallone, Ross, and Lepper (1985) have uncovered this same effect beyond the election setting and call it the "hostile media phenomenon."

Table 5. Concordance between Respondents and Their Media Intermediaries

	TV News	Newspapers
Media's candidate preference is:		
Concordant	9%	17%
Neutral	64%	60%
Discordant	27%	23%
Respondent-media correlation	−.31	−.18
Cases	386	376

NOTE.—This table pertains to voters and those exposed to each alter only. Cell entries are the percentages of voters who perceive the vote preferences of the particular medium to be the same as theirs, neither pro-Bush nor pro-Dukakis, or opposite to theirs, respectively. The correlations are τ_c's computed between a dichotomous measure of vote and perceptions that alters are pro-Bush, balanced or neutral, or pro-Dukakis, with the Dukakis alternative receiving the highest score on each measure. All coefficients significant at the.05 level are underlined.

These results shed new light on research on mass media effects. If the media supply little direct cuing for presidential preferences and, consequently, are perceived as balanced or misperceived as favoring the opposing candidate, then it is little wonder that their effects seem to be so modest and indirect.

Variations in ego-media concordance. Some understanding of the possibilities for media influence may be gained from an examination of the relationships between perceived ego-media agreement and the 10 exogenous variables employed previously. These data appear in table 6.

Overall, there is little association of demographic characteristics, political attentiveness, and partisan commitment with respondent perceptions of the match between media and their own preferences. Only four variables emerge as significant in the multivariate analysis. Two are measures of partisan commitment. As ideological commitment strengthens, so too does the perception that newspapers are favorable to one's own candidate. Conversely, as perceived differences between the candidates grow, television is more likely to be seen as favoring the opposing candidate. Perhaps the best explanation for this difference lies in the one palpable dissimilarity between the media in their treatment of the presidential campaign: newspapers endorse candidates. Ideologues are more attuned to this cue. In the absence of a comparable cue from television news, on the other hand, negative projection may take over.

Table 6. Characteristics of Concordance with Media Intermediaries

	TV News		Newspapers	
	r	b	*r*	b
Demographic characteristics:				
Race	.00	.01	−.12	−.23
Gender	−.03	−.02	.12	.13
Age	−.05	.00	−.04	.00
Education	−.02	.00	.14	.04
Political attentiveness:				
Campaign interest	−.18	−.09	−.09	−.13
Concern about outcome	−.15	−.03	.10	.09
Information	−.16	−.06	.03	−.04
Partisan commitment:				
Ideological strength	−.11	−.03	.16	.09
Party strength	−.11	−.03	−.05	−.04
Absolute candidate distance	−.19	−.002	.06	.00
Adjusted R^2		.05		.07
Cases	368		357	

NOTE.—This table pertains to voters and those exposed to each medium only. The entries in each column are Pearson correlations (*r*) and multivariate unstandardized regression coefficients (b) between each row variable and concordance (−1 = discordant, 0 = neither, +1 = concordant). Higher scores on personal characteristics are assigned to blacks, males, elderly, college educated, interested in the campaign, concerned about the outcome, well-informed, strongly ideological, strongly partisan, and candidate thermometer ratings that are farthest apart. Coefficients significant at the .05 level are underlined.

The only remaining regression coefficients of any magnitude are for campaign interest and education. Politically interested citizens are inclined toward negative projection for both television and newspapers, but only the relationship for newspapers is significant in the multivariate analysis. Better-educated newspaper readers, on the other hand, see more consonance than dissonance in newspaper favoritism. Their greater cognitive ability may enable them to discern the preferences communicated by an editorial endorsement.

The Joint Personal and Media Context of Intermediation

The content of messages from intermediaries colors the flow of political information to the individual. This content has been described separately for personal and media intermediaries. Table 7 now combines

Table 7. Partisan Concordance of Intermediation Contexts (Percentages)

Agreement with Personal Discussants	Agreement with the Mass Media: Discordant	Neutral	Concordant	Total
Discordant	3.7	4.7	4.7	13.2
Neutral	11.7	20.3	7.2	39.2
Concordant	19.4	21.3	6.9	47.6
Total	34.7	46.4	18.9	100.0
Cases	140	187	76	403

NOTE.—This table pertains to voters and those exposed to at least one personal and media alter only. The row and column categories are based on the agreement between the perceived vote preference/favoritism of each intermediary and the respondent's vote preference (discordant pairs = −1, concordant pairs = 1, and all voters with neutral intermediaries = 0). These scores then were summed for each type of intermediary, and the net result was trichomotized into discordant (<0), neutral (0), and concordant (>0). The sum of the cell entries may not exactly equal the column and row totals due to rounding error.

these personal and media environments to estimate the perceived partisan character of a fuller intermediation environment.[6]

When all the personal intermediaries are considered simultaneously, it is clear that personal intermediaries provided little challenge to Ohioans' presidential preferences in 1988. The combined TV and newspaper context, by contrast, was at once more neutral and more discordant. The media, in short, were almost three times as likely as personal networks to challenge the individual's preferences. To many, they provide about the only dissonant note in an otherwise supportive or neutral score.

Perhaps the most interesting figures appear in the interior cells of table 7 where personal and media contexts are combined. Almost half of the respondents perceived themselves in contexts containing some dissonance. And more than half of these faced a split context, with both discordant and concordant intermediaries, leaving 20.1 percent of all respondents in a basically discordant intermediation context. That so many find discordant preferences in their immediate intermediation environments suggests that intermediaries may be considerably more important as agents of change in preferences than might have been expected initially.

---

[6] This full intermediation environment is built up from summaries of the preferences/favoritism of each type of intermediary. A summary measure is constructed for personal and media intermediaries separately by summing concordant (= 1) and discordant preferences (= −1); neutral positions (= 0) have no impact on the total. Spouse is not counted for main discussant so as not to give it a double weight.

Even though the contrasting patterns for personal and media intermediaries may confound interpretation, a brief examination of the characteristics of people within the different joint contexts may shed some light on the political intermediation process. The means of each of the 10 characteristics employed previously were calculated within each of three contexts—discordant (combining wholly discordant and split = 44 percent), neutral (20 percent), and concordant (35 percent). Significant mean differences between these groups (using an *F*-test) appeared for only four of the variables—campaign interest, concern about the outcome, information level, and ideological strength.

The patterns for the three measures of attentiveness—campaign interest, concern about outcome, and information about politics—are similar, even in the multivariate analysis, and can be considered together. Discordant contexts contain the most attentive voters, while neutral ones harbor those who are least attentive. That attentiveness to politics and location in discordant intermediation environments go together probably reflects the greater politicization of attentives—that is, a greater tendency among them to discuss politics in a variety of environments. But this greater exposure to different viewpoints among attentives may be inconsequential because the large store of political information they possess can protect them from influence by dissonant environments (Converse 1966).

This pattern is reversed when ideological commitment is considered. Concordant intermediation contexts contain the most committed ideologues, on average. Perhaps they are the most determined to structure a supportive political environment. Or perhaps a one-sided environment promotes strength of conviction. Unfortunately, these alternative explanations cannot be tested with cross-sectional data.

Conclusion

What are the implications of these results for the flow of information from the often remote political world to the individual voter? The preceding analysis views the voter as nestled in a network of personal and media intermediaries that serve as conduits for political messages. The empirical results confirm that, by exposing themselves differentially to various intermediaries, people do influence their opportunities to receive consonant and dissonant information about the presidential candidates.

Measurement of perceived exposure to and the content of these intermediaries, however, yielded both the familiar and some surprises. That many people discuss politics with others in their immediate environment and that these discussions often are held between those with very similar political preferences are hardly novel findings. That relatives dominate as discussion partners also is not a new result, even if it is sometimes overlooked in personal network analysis. What is surprising is that the percentage in concordant personal contexts falls short of a majority because so many people find themselves in balanced or neutral discussion networks.

Also unexpected, because it has not been explored very much, is the extent to which political discussions in the workplace outnumber those in the neighborhood. These results suggest that the context for political discussion is not geographically delimited, as many contextual studies assume, but rather is defined both narrowly by family ties and broadly by opportunities for political interaction in the workplace. Contextual research needs to reexamine its assumption of the primacy of residential proximity (for the same point from another perspective, see Weatherford 1982).

By the same token, the finding of widespread exposure to television and newspapers is not unexpected. The mass media are universally credited with a wide reach in American society, as this study confirms. That regular usage was higher for newspapers than for television news, though, should give one pause in automatically attributing dominance to television in the intermediation process.

Most surprising of all was the direction of the partisan favoritism perceived in mass media intermediaries. The media were most commonly seen as balanced and neutral in the presidential race, as befits their self-image. Those who perceived them as favoring particular candidates, however, were more likely to view them as hostile than supportive. If these perceptions reflect reality, then the media are in a position to challenge the dispositions of many in their audience. If they distort reality, then an intriguing perceptual process seems to be at work in which many people, engaging in what we have called negative projection, mistake partisan neutrality in the media for favoritism toward the opposition. What this adds up to overall is a media that is not seen as lending support and reinforcement to the partisan proclivities of its audience. By its studied neutrality, perhaps, this media avoids directly challenging partisan views; but, by not supporting them, it may end up undermining them nonetheless.

The tendency toward homophyly with personal intermediaries, on the other hand, constrains their influence. There is plenty of room for reinforcement of political preferences but little room for change, because people do not often discuss politics with those who disagree with them. For some, of course, the possibility for reinforcement does not exist because they do not discuss politics at all, leaving them open to influence from the media. If the media too are neutral, the result might be withdrawal from presidential politics. Wouldn't it be ironic if balanced treatment of the presidential candidates in the media contributed to recent declines in the turnout of the American electorate?

Under some circumstances, though, personal intermediaries may play important roles in challenging preferences and perhaps even inducing partisan change. People who work outside of the home are more exposed to dissonant perspectives. This contact can overcome the inertial force of homophyly and play a significant role in electoral change. More exposed also are those who choose their main discussants from outside of the family, a tendency that is heightened by working outside the home. This documents for the political world the same phenomenon Granoveter (1973) has observed in the social world: it is the weakest (extrafamilial) ties that may exert the greatest influence.

Young adults are especially exposed to dissonant political views in their personal networks. This result has an interesting theoretical implication for the familiar finding of voting research that they are the most responsive of all age groups to prevailing period pressures in formulating their presidential preferences and even their partisanship. This greater youthful openness typically is explained by their weak partisan socialization (Beck 1974) and their lack of hardened political dispositions (Converse 1969, 1976). To these explanations must now be added the greater exposure of the young to challenging political views in their personal networks. Early adulthood, in short, is the life stage at which the individual and his or her political context are least harmonized.

Another distinctive electoral behavior of the young may be attributable to their more politically dissonant personal environments. Discord in political discussions seems likely to promote political hesitation and paralysis as individuals experience powerful cross-pressures in reaching an electoral decision. Following the familiar principles of cognitive dissonance, this discord leads some people to withdraw from the electoral arena altogether. Their greater likelihood to find themselves in politically dissonant environments, consequently, may contribute to the traditionally low turnout of the young.

This inquiry takes only preliminary steps in examining voters in their full intermediation context. Of course, much remains to be done. First, even if perceptions may define an individual's "real" world, it is important to gain a more objective assessment of the flow of information from intermediaries. Intermediation also is a dynamic process, often involving a reciprocal exchange of information. A second future task is to study this process using a longitudinal design. Finally, it is important to address how intermediation differs across space and time, for this may lead to greater understanding of political stability and change. A third challenge, therefore, is to study intermediation in other places and other times beyond Ohio in 1988. I hope that this initial inquiry will stimulate more research on the voter's full intermediation context that can fulfill the promise of those landmark studies of intermediation conducted half a century ago.

Appendix

Survey Questions

MEDIA INTERMEDIARIES

Television news. Exposure: "In the average week, how often do you watch the national evening news—never, once in a while, most days, or every day?" *Preference:* "On the whole, do you feel that the national evening news you watched favored Michael Dukakis, favored George Bush, or were they pretty well balanced between the two?"

Newspapers. Exposure: "In an average week, how often do you read a daily newspaper—never, once in a while, most days, or every day?" *Preference:* "What newspaper do you read the most? Do you feel (it) on the whole favored Dukakis, Bush, or was it pretty well balanced between them?"

PERSONAL INTERMEDIARIES

General discussion. "Some people don't talk about politics at all. What about you? Have you talked about the 1988 presidential campaign, the candidates, and the issues with anyone, such as a friend, neighbor, co-worker, your spouse, or other relative?" [Respondents answering "No" were not asked the following questions.]

Spouse. Exposure: "Have you talked about the 1988 presidential campaign, the candidates, and the issues with your spouse, or are you not married? [If yes:] How often was that—a few times or many times?" *Preference:* "In the election this year, did your spouse favor the Democratic candidate—Michael Dukakis—the Republican candidate—George Bush—or neither candidate?"

Neighbors. Exposure: "Have you talked about the 1988 presidential campaign, the candidates, and the issues with any of your neighbors?" [If yes:] "How often was that—a few times or many times?" *Preference:* [If talked to more than one:] "In the election this year, were the neighbors you talked to all for Dukakis, mostly for Dukakis, evenly split between Dukakis and Bush, mostly for Bush, or all for Bush?" [If talked to one:] "In the election this year, was the neighbor you talked to for Dukakis or Bush?"

Co-workers. Exposure: "Do you currently work outside of your home?" [If yes:] "Have you talked about the 1988 presidential campaign, the candidates, and the issues with any people where you work?" [If yes:] "How often was that—a few times, or many times?" *Preference:* [If talked to more than one:] "In the election this year, were they all for Dukakis, mostly for Dukakis, evenly split between Dukakis and Bush, mostly for Bush, or all for Bush?" [If talked to one:] "In the election this year, was the co-worker you talked to for Dukakis or Bush?"

Main discussant. Exposure: "Who did you talk with the most about the presidential campaign? Did you talk with this person about the presidential campaign a few times or many times?" *Preference:* "In the election this year, did this person favor Dukakis, Bush, or neither?"

DEMOGRAPHIC CHARACTERISTICS

Race. "What is your race?"

Education. "What is the highest grade of school or year of college you completed?"

Age. "In what year were you born?"

Gender. [Interviewer determination]

POLITICAL ATTENTIVENESS

Campaign interest. "Some people don't pay much attention to political campaigns. How about you? Would you say that you were very much interested, somewhat interested, or not much interested in following the political campaign this year?"

Concern about outcome. "Generally speaking, would you say that you personally cared a good deal which candidate won the election for President, that you cared somewhat, or that you did not care very much which candidate won?"

Political information. Interviewer determination: "R's general level of information about politics and public affairs seemed—very high, fairly high, average, fairly low, very low?"

POLITICAL COMMITMENT

Ideological commitment. "When it comes to politics in general, do you usually think of yourself as a conservative, as a liberal, as middle-of-the-road, or don't you think of yourself along liberal-conservative lines?" [If conservative or liberal:] "Do you think of yourself as a strong or not very strong [conservative/liberal]?"

Partisan commitment. "Generally speaking, do you usually think of yourself as a Republican, a Democrat, an independent, or what?" [If Democrat or Republican:] "Would you call yourself a strong [Democrat/Republican] or a not very strong [Democrat/Republican]?" [If independent, no preference, or other:] "Do you think of yourself as closer to the Republican party or to the Democratic party?"

Absolute candidate distance. Based on evaluations of Bush and Dukakis asked in the following way: "I'd like to get your feelings toward some of our political leaders and other people who are in the news these days. I'll read the name of a person and I'd like you to rate that person using something we call the feeling thermometer. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the person. Ratings between 0 degrees and 50 degrees mean that you don't feel favorable toward the person and that you don't care too much for that person. You would rate the person at the 50 degree mark if you don't feel particularly warm or cold toward the person. Our first person is ———. How would you rate him using the thermometer?"

References

Beck, Paul Allen. 1974. "A Socialization Theory of Partisan Realignment." In *The Politics of Future Citizens*, ed. Richard G. Niemi, pp. 199–219. San Francisco: Jossey-Bass.

Berelson, Bernard, Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting*. Chicago: University of Chicago Press.

Brickell, Bettina, Robert Huckfeldt, and John Sprague. 1988. "Gender Effects on Political Discussion: The Political Networks of Men and Women." Paper presented at the annual meeting of the American Political Science Association, Washington, DC.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

Chaffee, Steven H. 1982. "Mass Media and Interpersonal Channels: Competitive, Convergent, or Complementary?" In *Inter/Media: Interpersonal Communication in a Media World*, ed. Gary Gumpert and Robert Cathcart, pp. 57–77. New York: Oxford University Press.

Converse, Philip E. 1966. "Information Flow and the Stability of Partisan Attitudes." In *Elections and the Political Order*, ed. Angus Campbell, Philip E. Converse, Warren E. Miller, and Donald E. Stokes, pp. 136–57. New York: Wiley.

———. 1969. "Of Time and Partisan Stability." *Comparative Political Studies* 2:139–71.

———. 1976. *The Dynamics of Party Support*. Beverly Hills, CA: Sage.

Erbring, Lutz, Edie N. Goldenberg, and Arthur H. Miller. 1980. "Front-Page News and Real-World Cues: A New Look at Agenda-setting by the Media." *American Journal of Political Science* 24:16–49.

Eulau, Heinz. 1980. "The Columbia Studies of Personal Influence." *Social Science History* 4:207–28.

Fiorina, Morris. 1981. *Retrospective Voting in American National Elections*. New Haven, CT: Yale University Press.

Frey, James H. 1989. *Survey Research by Telephone*, pp. 99–100. Newbury Park, CA: Sage.

Graber, Doris. 1980. *Mass Media and American Politics*. Washington, DC: Congressional Quarterly.

Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360–80.

Huckfeldt, Robert, and John Sprague. 1987. "Networks in Context: The Social Flow of Political Information." *American Political Science Review* 81:1197–1216.

Iyengar, Shanto. 1987. "Television News and Citizens' Explanations of National Affairs." *American Political Science Review* 81:815–31.

Iyengar, Shanto, and Donald R. Kinder. 1987. *News That Matters*. Chicago: University of Chicago Press.

Katz, Elihu, and Paul F. Lazarsfeld. 1955. *Personal Influence*. Glencoe, IL: Free Press.

Kinder, Donald R., and David O. Sears. 1985. "Public Opinion and Political Behavior." In *The Handbook of Social Psychology*, ed. Gardner Lindzey and Elliot Aronson, 2:659–741. New York: Random House.

Klapper, Joseph. 1960. *The Effects of Mass Communications*. New York: Free Press.

Laumann, Edward O. 1973. *Bonds of Pluralism: The Form and Substance of Urban Social Networks*. New York: Wiley.

Lazarsfeld, Paul, Bernard Berelson, and Hazel Gaudet. 1948. *The People's Choice*. New York: Columbia University Press.

Niemi, Richard G. 1974. *How Family Members Perceive Each Other*. New Haven, CT: Yale University Press.

Nimmo, Dan, and James E. Combs. 1983. *Mediated Political Realities*. New York: Longman.

Patterson, Thomas E. 1980. *The Mass Media Election: How Americans Choose Their President*. New York: Praeger.

Patterson, Thomas E., and Robert D. McClure. 1976. *The Unseeing Eye*. New York: Putnam's.

Ranney, Austin. 1983. *Channels of Power*. New York: Basic.

Roberts, Donald F., and Nathan Maccoby. 1985. "Effects of Mass Communication." In *The Handbook of Social Psychology*, ed. Gardner Lindzey and Elliot Aronson, vol. 2, pp. 539–98. New York: Random House.

Robinson, Michael J., and Margaret A. Sheehan. 1983. *Over the Wire and on TV: CBS and UPI in Campaign '80*. New York: Russell Sage.

Sartori, Giovani. 1969. "The Sociology of Parties: A Critical View." In *Party Systems, Party Organizations and the Politics of the New Masses*, ed. Otto Stammer. Berlin: Institute für Politische Wissenschaft, Freien Universitat.

Sheingold, Carl A. 1973. "Social Networks and Voting: The Resurrection of a Research Agenda." *American Sociological Review* 38:712–20.

Sudman, Seymour. 1983. "Applied Sampling." In *Handbook of Survey Research*, ed. Peter H. Rossi, James D. Wright, and Andy B. Anderson, pp. 177–78. New York: Academic Press.

Vallone, Robert P., Lee Ross, and Mark R. Lepper. 1985. "The Hostile Media Phenomenon: Biased Perception and Perceptions of Media Bias in Coverage of the Beirut Massacre." *Journal of Personality and Social Psychology* 49:577–85.

Weatherford, M. Stephen. 1982. "Measurement Problems in Contextual Analysis: On Statistical Assumptions and Social Processes." *Political Methodology* 8:61–90.

Zaller, John. 1985. "Pretesting Information Items on the 1986 Pilot Survey." Report submitted to the National Election Studies Board of Overseers.

HARDLY EVER OR CONSTANTLY?
GROUP COMPARISONS USING VAGUE QUANTIFIERS

NORA CATE SCHAEFFER

Abstract Investigations of vague quantifiers—phrases like "very often," which express frequency in relative terms—have left the consequences of analyzing absolute or relative responses largely unexplored. This paper examines whether the choice of relative or absolute frequencies affects comparisons by race, sex, education, and age and the meaning of relative-frequency and relative-intensity phrases for these groups. Items analyzed are the frequency of excitement and boredom (using a 1975 national probability sample of households with quotas, $N = 1{,}172$). Conclusions about race differences in excitement and boredom differ for absolute and relative frequencies. There are significant differences by race, education, and age in the meaning of phrases conveying relative frequency: such phrases stand for higher absolute frequencies for whites, for better educated, and for younger respondents.

Introduction

In the movie *Annie Hall* there is a scene with a split screen. On one side Alvie Singer talks to his psychiatrist; on the other side Annie Hall talks to hers. Alvie's therapist asks him, "How often do you sleep together?" and Alvie replies, "Hardly ever, maybe three times a week." Annie's therapist asks her, "Do you have sex often?" Annie replies, "Constantly. I'd say three times a week." That the same abso-

NORA CATE SCHAEFFER is associate professor of sociology at the University of Wisconsin—Madison. This research was supported in part by a grant from the graduate school of the University of Wisconsin—Madison. The author thanks Charles N. Halaby, Rob Mare, Judith A. Seltzer, Elizabeth Thomson, Roger Tourangeau, and two anonymous referees for their comments; Theresa Hammond for watching *Annie Hall;* Patrick Bova of NORC for help in obtaining survey and data file documentation; and Hong-wen Charng, Kay Tuschen, and Elizabeth Uhr for their assistance in preparing this paper. An earlier version of this paper was presented at the annual meeting of the American Association for Public Opinion Research, Toronto, May 1988.

Public Opinion Quarterly Volume 55:395–423 © 1991 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/91/5503-0004/$2.50

lute frequency can be reported in very different ways is funny in a movie, but is a source of concern for researchers. This paper investigates how the meaning of phrases expressing relative frequency varies across groups and how these variations affect group comparisons.

Response categories for questions in experiments or surveys are often labeled with vague quantifiers or intensifiers (Bradburn and Miles 1979), phrases like "very often" or "very much" that express relative frequency or intensity. Most vague quantifiers or intensifiers imply a comparison to an implicit standard, as in deciding whether an event occurs "often," and therefore elicit relative responses. Previous research has focused on identifying components of the meaning of these phrases (Mosier 1941), identifying factors that modify their meanings (e.g., Chase 1969; Parducci 1982), and developing scales that express distances between them. Studies have scaled phrases indicating relative intensity (Jones and Thurstone 1955; Lodge et al. 1976; Mosier 1941; Spector 1976), adverbs used as intensifiers (Cliff 1959), probabilities (Lichtenstein and Newman 1967), and phrases expressing relative frequency or amounts (e.g., all, some, etc.) (Bass 1968; Bass, Cascio, and O'Conner 1974; Dennis, Dennis, and Jobe 1975; Hakel 1969; Newstead and Collis 1987; Pace and Friedlander 1982; Pepper 1981; Schriesheim and Schriesheim 1978; Simpson 1944; Spector 1976). The absolute frequency implied by a relative frequency depends on how often the event actually occurs; "often" is less often for an unusual event, like an earthquake, than for an event that occurs frequently, such as gunfire in Westerns (Pepper and Prytulak 1974). Feelings also affect judgments of relative frequency, so that "often" may mean fewer unpleasant than pleasant events. Those who dislike eating alone, for example, may feel that "often" describes eating alone three times a week, whereas those who like eating alone would not call that frequency "often" (Goocher 1965).

Frequency is probably the most common response dimension in questions about events or behaviors (Schaeffer [1989] describes alternatives). A frequency response dimension can be implemented by asking for absolute frequencies, either raw or grouped, or by requesting relative frequencies.[1] The impact on analyses of using relative versus absolute frequency categories is largely unexplored. (I refer to raw counts of events or behaviors as absolute responses, although "absolute" here does not imply that such frequency reports are accurate.) Bradburn and Miles (1979) scored the categories of relative-frequency items using equal intervals (as is usually done) and using the average absolute frequency implied by the category. The two scoring systems yielded similar correlations between the relative-frequency items and a set of criterion variables. Based on this comparison, Bradburn and Miles found little advantage in the absolute frequencies and cautioned that respondents find them more difficult to provide. After evaluating errors in reports of absolute frequencies, Bradburn and Danis (1984) conclude that respondents are better at making relative than absolute frequency judgments. If the frequency of an event or feeling about an event varies by group, however, the meaning of phrases expressing relative frequency may also vary, and this may affect group comparisons. For example, if, on average, men experience a symptom less frequently than women (or are more embarrassed by the symptom), "often" could mean a lower frequency for men. When the meaning of vague quantifiers differs by group, relative and absolute responses may lead to different conclusions.

---

[1] A common method for obtaining relative frequencies is to offer categories such as "very often, pretty often, not too often, never," but sometimes variants with a more explicit standard such as "more than usual, less than usual, or about the same as usual" are used. Some frequency response categories labeled with vague quantifiers imply proportional judgments. The categories "all of the time, most of the time, some of the time, never," e.g., imply that the number of times an event occurred can be compared to the number of opportunities for it to have occurred. Such proportional judgments are not discussed here.

I use secondary analyses of a national study to consider whether conclusions about between-group differences are the same when absolute and relative responses are examined. The analysis then examines between-group differences in the meaning of relative-frequency categories. This analysis "scales" the relative-frequency phrases using reports of absolute frequencies. I present four common group comparisons: race, sex, education, and age. Because the existing literature does not systematically distinguish between relative and absolute frequencies, the analysis is exploratory. I do not develop specific hypotheses about the relationship between the group variables and the type of responses. The distinction between absolute and relative responses and the potential impact of this choice on research conclusions are relevant to a wide range of questions used in both surveys and experiments.[2]

Data

Data on the meaning of relative frequencies are provided by an experiment conducted by National Opinion Research Center (Bradburn, Sudman, and Associates 1979). Personal interviews were obtained in 1975 with 1,172 adults selected in a national probability sample with quotas.

---

[2] The question is also relevant, but less tractable, for relative intensity categories (see Schaeffer and Bradburn 1989, pp. 409–11).

Deviations from this number in the tables are due to item nonresponse and to the exclusion of races other than blacks and whites in race comparisons. Respondents were asked to estimate relative frequencies for one positive event (being excited) and one negative event (being bored) in the Affect Balance Scale (ABS) (Bradburn 1969) as follows: "We are interested in the way people are feeling these days. How often do you feel . . . particularly excited or interested in something? Very often, pretty often, not too often, or never? . . . bored?" (Leisure Activity Study of the Response Effects in Surveys project [1980]). For each of these two items, respondents who selected a relative frequency other than "never" were then asked, "About how many times a week or a month did you mean?" Frequency estimates were converted to the number of times per month during coding.[3]

GROUP COMPARISONS OF EXCITEMENT AND BOREDOM

Does the decision to use relative or absolute frequencies determine whether between-group differences appear? Table 1 displays group comparisons by race, sex, education, and age for the frequency of being excited and bored. For each item, the results based on relative frequencies are followed by the results based on raw absolute frequencies and their logs. Because this analysis is exploratory, in a few instances the discussion describes observed relationships as well as their significance (alpha = .05, under the assumption of simple random sampling).

The results in table 1 indicate that there is no significant difference between blacks and whites in the relative frequency of excitement ($p = .98$). When absolute frequencies are examined, however, excitement appears to be more frequent for whites than blacks; the difference is not significant when raw absolute frequencies are tested ($p = .17$), but it is when log absolute frequencies are examined ($p < .05$). The conclusion about race differences in the frequency of boredom is not the same for relative and absolute frequencies: the relative frequency of boredom is significantly greater for blacks than whites ($p < .01$), but the absolute frequencies suggest there is no group difference. There is no significant sex difference in excitement or boredom using either type of response. For education, the relative and absolute responses agree that those with more education feel excited more often ($p < .01$

---

[3] Not asking for the absolute frequency of zero was understandable but unfortunate because respondents may mean a frequency greater than zero when they say "never" (Pohl 1981, p. 239). Because there is a much higher consensus on the meaning of "never" than of many other vague quantifiers (Hakel 1969; Simpson 1944), however, the design probably only slightly affects the strength of the group-by-category interactions examined later. Analyses omitting these cases gave results almost identical to those reported.

Table 1. Relative and Absolute Frequencies of Being Excited and Being Bored, by Race, Sex, Education, and Age

	Race		Sex		Education			Age			
Frequency Item	Black	White	Male	Female	<12	12	>12	<31	31–45	46–60	>60
Excited, relative frequencies:											
Never	3.2	2.1	1.8	2.7	4.8	.7	1.3	.3	1.3	2.7	6.2
Not too often	31.0	30.4	35.3	27.1	40.7	30.2	22.5	25.6	27.9	33.9	40.5
Pretty often	40.6	45.4	39.9	48.1	33.2	47.3	52.1	52.1	50.2	39.3	29.1
Very often	25.2	22.1	22.9	22.1	21.3	21.8	24.1	22.0	20.6	24.1	24.2
N	155	973	546	621	376	404	382	359	315	257	227
Mean	1.88	1.88	1.84	1.90	1.71	1.90	1.99	1.96	1.90	1.85	1.71
S.D.	.82	.77	.79	.77	.85	.74	.72	.70	.73	.82	.90
	$F_{1,\,1,126} = .00$, $p = .98$		$F_{1,\,1,165} = 1.52$, $p = .22$		$F_{2,\,1,159} = 12.99$, $p < .01$			$F_{3,\,1,154} = 4.85$, $p < .01$			
Excited, raw absolute frequencies:											
Mean	10.53	12.03	12.10	11.43	9.33	11.68	14.09	13.65	12.64	9.95	9.05
S.D.	11.22	12.38	13.31	11.09	9.76	11.69	14.18	12.33	12.86	10.28	11.14
	$F_{1,\,1,056} = 1.87$, $p = .17$		$F_{1,\,1,094} = .81$, $p = .37$		$F_{2,\,1,089} = 13.76$, $p < .01$			$F_{3,\,1,084} = 8.87$, $p < .01$			

Table I. (Continued)

	Race		Sex		Education			Age			
Frequency Item	Black	White	Male	Female	<12	12	>12	<31	31–45	46–60	>60
Excited, log absolute frequencies:											
Mean	.87	.95	.94	.93	.83	.96	1.01	1.03	.97	.88	.80
S.D.	.42	.39	.39	.40	.42	.36	.40	.36	.40	.38	.43
	$F_{1, 1,056} = 4.85$, $p = .03$		$F_{1, 1,094} = .46$, $p = .50$		$F_{2, 1,089} = 20.71$, $p < .01$			$F_{3, 1,084} = 18.17$, $p < .01$			
Bored, relative frequencies:											
Never	24.2	31.9	30.8	30.3	30.1	30.7	30.5	20.0	32.0	33.5	41.3
Not too often	46.4	49.9	48.1	50.7	46.1	51.5	50.9	55.0	51.3	43.6	45.3
Pretty often	11.8	11.4	11.7	10.6	12.0	9.4	12.3	13.6	7.9	14.4	8.0
Very often	17.6	6.8	9.4	8.4	11.7	8.4	6.3	11.4	8.9	8.6	5.3
N	153	974	545	621	375	404	383	360	316	257	225
Mean	1.23	.93	1.00	.97	1.05	.96	.94	1.16	.94	.98	.77
S.D.	1.01	.84	.90	.86	.94	.86	.82	.88	.87	.91	.81
	$F_{1, 1,125} = 15.86$, $p < .01$		$F_{1, 1,164} = .24$, $p = .62$		$F_{2, 1,159} = 1.82$, $p = .16$			$F_{3, 1,154} = 9.87$, $p < .01$			

Bored, raw absolute frequencies:											
Mean	5.45	4.85	5.39	4.64	4.52	5.09	5.34	6.90	4.30	4.43	3.53
S.D.	7.92	7.84	8.61	7.15	6.98	8.00	8.53	8.73	7.79	7.03	6.91
	$F_{1,\ 1,093} = .73$, $p = .39$		$F_{1,\ 1,132} = 2.59$, $p = .11$		$F_{2,\ 1,127} = 1.04$, $p = .35$			$F_{3,\ 1,122} = 10.89$, $p < .01$			
Bored, log absolute frequencies:											
Mean	.56	.50	.53	.50	.50	.52	.53	.66	.46	.48	.39
S.D.	.46	.46	.47	.45	.44	.46	.47	.46	.44	.45	.43
	$F_{1,\ 1,093} = 2.00$, $p = .16$		$F_{1,\ 1,132} = .67$, $p = .41$		$F_{2,\ 1,127} = .37$, $p = .69$			$F_{3,\ 1,122} = 20.53$, $p < .01$			

NOTE.—Total sample size is 1,172 including all races and ignoring item nonresponse. Relative frequency entries are percentages. Relative frequency categories were scored 0 to 3 to calculate means. Absolute frequencies are described in the Appendix.

for all). The three tests also agree that there is no significant difference among education groups in the frequency of boredom, although the direction of the observed relationship with education differs for the relative and absolute frequencies. For age, the relative and absolute frequencies both indicate that those who are younger are excited and bored significantly more often than those who are older ($p < .01$ for all).[4]

These results suggest that the choice of relative or absolute frequencies can affect conclusions about the significance of group differences or the observed direction of group differences that are not significant. The significance of results is affected only for race, and only consistently for the logs of the absolute frequencies, but it is affected for both excitement and boredom. Regardless of whether the two metrics lead to different conclusions, however, groups may differ in the absolute frequency they mean when they use a relative-frequency phrase.

GROUP DIFFERENCES IN THE MEANING OF RELATIVE FREQUENCIES

The literature reviewed earlier indicates that, on average, "often" means more often for a common (or positive) event than for a rare (or negative) event. This literature predicts that the relative-frequency categories will indicate a higher absolute frequency for excitement than for boredom (comparing events) and for the group with the greater mean absolute frequency overall (comparing groups). These patterns should be reflected in the group means for each of the relative frequencies. The average absolute frequency implied by a relative-frequency phrase for a group can be thought of as the average meaning of the phrase for the group. Thus, the group mean is a kind of scale value for the phrase when scale values are based on absolute frequency reports.[5] Before examining group differences in the absolute frequency implied by relative-frequency phrases, however, some features of the absolute frequencies should be described.

---

[4] The comparison in table 1 is presented in the exploratory spirit of this analysis and presents the absolute and relative frequency comparisons as though each were from a "one shot" survey. Comparing results for relative and absolute frequencies is difficult because the models are not nested—the metrics are inherently incomparable—and the questions obtaining the answers are not independent. When the absolute frequencies are transformed, the comparison becomes even more intractable. Ignoring these issues, one could analyze the data presented in table 1 as a repeated measures analysis of variance (either separately for excitement and boredom or considering the items together), although the results must be interpreted cautiously because of the differences in metric (relative vs. absolute frequency) and the effect of metric on variances. For the raw absolute frequency of excitement, the group differences that appear using the two metrics differ significantly ($p < .05$) for education and age; none is significant when log absolute frequencies are analyzed. For the absolute frequency of boredom, the group difference obtained using the two metrics differs significantly for age; using the log absolute frequencies, this difference is significant for race, education, and age.

[5] Although my analysis focuses on means, issues concerning the variances also require attention. For example, the variation in meaning of vague probability terms appears to

Within each relative-frequency category, the lowest absolute frequency reported is one, answers are clustered at low values—particularly for "not too often" and "pretty often"—and there is a long positive tail. (Appendix table A1 summarizes the distribution of absolute frequency within relative-frequency category.) Moreover, within five of the six relative-frequency categories, there is at least one very high outlier. The incidence of high values increases as relative frequency increases.[6] Finally, when expressed in the original metric, the mean of the log frequencies (the geometric mean) reflects the influence of the frequently selected low values better than does the arithmetic mean. For "not too often" and "pretty often" the geometric mean is closer to the mode and median than is the arithmetic mean.

Table 2 presents the means and standard deviations of raw and log absolute frequency by relative-frequency category and group. The data illustrate the consequences of the features just described. The variance of the raw absolute frequencies tends to increase as the relative frequency increases and as the mean of the raw absolute frequencies increases. (Considering each of the groups by relative-frequency category comparisons, the ratio of the largest to the smallest variance ranges from approximately 5 to 16.) For the 14 comparisons between the means of adjacent groups, the absolute group difference is largest for "very often" in 10 cases, and the pattern is clearest for boredom. Finally, as Bradburn and Miles's (1979) published data suggest, the means for "not too often" are higher for excitement than boredom, but that is not true for either "pretty often" or "very often."

An analysis of these data must consider their general structural features. The reports of absolute frequencies refer to the prior relative-frequency report, and absolute frequency reports cannot be treated as

---

increase as the probability approaches .5 and decrease as the phrase indicates greater or lesser likelihood (Wallsten et al. 1986, p. 354). In the data analyzed here, the smallest reported absolute frequency is relatively similar for each of the relative frequencies, but the largest absolute frequency increases with the relative frequency. In such cases, the variance of the absolute frequencies will also be largest for the relative frequency phrase with the highest mean absolute frequency.

[6] In these data, "high outliers" are numbers equal to 97 or greater. These were assigned the value 97 in the original coding of the data. For the race comparison, outliers occur only among whites for both excitement and boredom. Both males and females report outliers for both excitement and boredom. There are no outliers for those with less than a high school education, and outliers are less frequent for younger than older respondents. With the exceptions just noted, the distribution of outliers appears haphazard for sex, education, and age. Detailed comparisons indicate that the only consequence of recoding the outliers (see the App.) is to reduce means trivially and slightly to reduce the variances for raw absolute frequencies. Effects on means and variances of log absolute frequencies are trivially small. There is no effect on conclusions from significance tests.

Table 2. Means and Standard Deviations for Absolute Frequency of Being Excited and Being Bored by Relative Frequency and Race, Sex, Education, and Age

	Race		Sex		Education			Age			
Frequency Item	Black	White	Male	Female	<12	12	>12	<31	31–35	46–60	>60
Excited, frequencies:											
Not too often:											
Mean	6.22	6.63	6.76	6.28	6.27	7.57	5.41	6.84	6.98	6.73	5.56
S.D.	5.61	7.78	7.25	7.73	7.18	8.93	5.10	6.19	9.04	6.47	7.98
Pretty often:											
Mean	12.58	13.00	13.23	12.71	11.37	12.24	14.65	14.81	12.14	10.45	12.17
S.D.	13.37	11.86	13.00	11.23	10.24	11.42	13.43	12.72	9.87	9.56	13.63
Very often:											
Mean	14.15	18.16	18.53	16.59	13.76	16.60	21.49	18.96	21.84	14.54	13.35
S.D.	10.98	14.96	16.96	11.51	10.91	13.55	16.82	13.35	17.65	13.46	10.40
Excited, logs of frequencies:											
Not too often:											
Mean	.61	.63	.64	.60	.58	.69	.57	.67	.61	.68	.53
S.D.	.43	.40	.41	.40	.42	.39	.38	.39	.45	.36	.40
Pretty often:											
Mean	.88	.96	.95	.94	.87	.94	1.00	1.02	.94	.84	.88
S.D.	.48	.38	.40	.40	.43	.37	.39	.38	.38	.41	.43

Very often:											
Mean	.95	1.11	1.11	1.07	.96	1.08	1.20	1.18	1.20	.96	.96
S.D.	.48	.39	.39	.42	.45	.39	.36	.31	.39	.47	.42
Bored, frequencies:											
Not too often:											
Mean	4.54	3.99	4.17	3.94	3.74	4.40	3.96	4.46	3.54	3.82	4.26
S.D.	6.21	4.11	4.49	4.31	3.64	5.44	3,72	4.26	3.91	3.77	5.82
Pretty often:											
Mean	8.00	14.73	15.56	11.92	10.51	14.30	16.21	17.29	11.75	10.09	13.83
S.D.	6.20	10.98	11.78	9.13	9.50	11.84	10.10	11.19	9.74	8.73	11.67
Very often:											
Mean	13.77	18.29	17.35	16.65	13.90	18.06	20.75	18.15	18.15	15.59	11.67
S.D.	10.86	11.25	12.24	9.93	10.19	8.16	14.65	7.92	14.64	10.94	11.20
Bored, logs of frequencies:											
Not too often:											
Mean	.46	.46	.46	.45	.44	.47	.46	.52	.40	.43	.44
S.D.	.38	.34	.35	.34	.33	.37	.34	.33	.33	.36	.37
Pretty often:											
Mean	.74	1.01	1.01	.93	.81	1.01	1.10	1.13	.91	.84	.88
S.D.	.42	.41	.46	.38	.47	.37	.36	.33	.41	.39	.57
Very often:											
Mean	.94	1.16	1.11	1.09	.98	1.20	1.16	1.20	1.09	1.01	.88
S.D.	.50	.34	.38	.41	.44	.26	.44	.27	.43	.49	.44

NOTE.—The number of cases in each cell can be computed from table 1.

true frequencies because they have systematic errors compared with true absolute frequencies (Bradburn and Danis 1984). An analysis also confronts the problems of heteroscedasticity and of how to represent the "central tendency" of such skewed distributions. An analysis using raw frequencies may be particularly sensitive to high outliers. Using the geometric mean to summarize the "central tendency" of the distribution of absolute frequencies is compatible with practice in psychophysical and other methods of scaling. In such studies, distributions of the kind found here are common, and the ability to reproduce the group total from the arithmetic mean is not important. Using the median or geometric mean to represent a distribution of responses reduces the influence of high outliers (see, e.g., Lodge [1981] and Torgerson [1958] on scaling and Rubin [1984] on transformations). Although a previous analysis of these data analyzed the raw frequencies (Bradburn and Miles 1979), these considerations suggest analyzing logs, particularly for making group comparisons.

Analyzing logs of the absolute frequencies has several consequences. As the summary statistics in table 2 indicate, analyzing log absolute frequencies reduces heteroscedasticity. (For log absolute frequencies the ratio of the largest to the smallest variance in each group by category comparison ranges from approximately 1 to 3.) Most of the group differences calculated using the geometric means are smaller than those calculated using the arithmetic means. This is because the geometric mean reduces the influence of high values on group comparisons. There are some problems with analyzing logs, however. One could argue that instead of enabling one to arrive at a more useful representation of the middle of a skewed distribution, the transformation treats differences in variances between groups as artifactual and ignores the substantive issue the variances present. Also, inferences on the transformed data do not refer to the original metric. Comparing results obtained using raw and transformed scores is not straightforward, because the implicit models differ, and they are not nested. Finally, one must resort to ad hoc methods for treating values of zero—such as implicitly occur here for "never"—when moving to a log scale.

In examining group differences in the average meaning of these phrases, however, it seems reasonable to give the large proportion of answers bunched at low frequencies more weight and to control the influence of high outliers on the means and variances of the raw absolute frequencies. For these reasons, and because reproducing the group total from the arithmetic mean is not important here, the analysis of absolute frequency by relative-frequency category and social group uses the transformed frequencies. (I note instances in which significance tests differ for raw frequencies.) A theoretical justification for analyzing logs might argue that respondents use relative-frequency phrases to reflect percentage, rather than raw, change in absolute frequencies. If respondents consider percentage change, the mean absolute frequency implied by low relative frequencies should show relatively large differences between events of different frequencies. Greater relative frequencies (e.g., "very often") might then show constant percentage increases in the absolute frequency implied.[7]

To determine whether the meaning of the phrases describing relative frequency varies by social group, I estimated regression equations predicting log absolute frequency from relative-frequency category and group membership as defined in table 2. This approach provides straightforward comparisons to the literature cited earlier on the meaning of relative-frequency words. Dummy variables for the relative-frequency categories, with "never" as the reference category, are used to predict absolute frequency. (I refer to the set of these dummy variables as the "category variables.") Because the equations predicting excitement and boredom use the same observations, their disturbances are likely to be correlated. In this case, the OLS estimator is not efficient, so the two equations are estimated simultaneously as a system of "seemingly unrelated regressions" (Felmlee and Hargens 1988; Greene 1988; Maddala 1977).[8] The model also provides straightforward tests of whether the effects of independent variables are the same for excitement and boredom and pooled tests of the significance of effects estimated in common across the two equations.

There are at least two issues to explore in comparing excitement and boredom. First, Bradburn and Miles's (1979) published data and table 2 suggest that "not too often" means a different frequency for excitement than for boredom, but the other relative-frequency phrases may not. Second, even if groups use relative frequencies differently, there is no reason to expect that pattern to vary for excitement and boredom. Before examining differences between excitement and boredom, I fit a model that constrained the intercept to be the same for excitement and boredom but left the other coefficients free. (These results are not shown separately; they differ little from those for the final model presented below.) The results show substantial and significant increases in the log absolute frequency associated with each of the relative-frequency categories; these effects are similar for excitement and boredom. Given the structure of the data, one would expect no group differences in the value of the intercept, and none appear. The group difference in the absolute frequency implied by the relative frequencies is almost always in the same direction for excitement and boredom. None of the group differences in the meaning of "not too often" is significant, with a single exception in the age comparison. But the group differences in the log absolute frequency implied by "pretty often" and "very often" are consistently significant (or borderline) for boredom: the absolute frequency is greater for whites, the better educated, and the young. There are no significant sex differences.[9]

---

[7] The use of relative frequencies to express percentage change is compatible with the finding of previous research that "often" implies a higher absolute frequency for a common than for a rare event. Such an argument would also be compatible with some findings in psychophysical scaling using magnitude estimation, but I do not know of any theories concerning frequency reports that suggest preferring either raw or log frequencies for this analysis.

[8] These models require that both dependent variables be observed for every respondent. A large proportion of the sample said that they were never bored, but most of these reported being excited. Omitting cases in which respondents said they were never excited or never bored would reduce the sample size considerably by excluding many cases in which the respondent reported being excited and would introduce a kind of selectivity. Analysis indicates that treating the cases that were never bored or excited differently would not change the conclusions reported here.

If differences between excitement and boredom are not significant, information contained in the two variables can be pooled to improve estimates of category and group differences. To explore whether the apparent differences between excitement and boredom are significant, I fit a series of models imposing equality constraints across the two equations. Imposing these constraints is compatible with the results for the unconstrained model. (The only exception is the model for sex, which is discussed separately below.) Panel A of table 3 presents fit statistics for models that impose cross-equation constraints. For example, model A requires that all effects be the same for excitement and boredom, and model B constrains all effects to be the same for excitement and boredom except the coefficient for "not too often." Panel B shows contrasts between these models. The contrast between model A and model B shows that releasing the constraint for the category "not too often" results in a significant improvement in fit. Releasing cross-equation constraints for the group difference in "not too often" (model C), the other group differences (model D), or the other relative-frequency categories (model E) or the group variables (model F) does not improve the fit, and so model B is accepted.

Table 4 presents the coefficients for model B. Only the estimate for the absolute frequency implied by "not too often" differs for excitement and boredom, so the table presents pooled estimates for the other coefficients. There are significant group differences in the meaning of "pretty often" and "very often": those with more education mean a greater absolute frequency by "pretty often" and "very often" than do those with less education. Similarly, those who are over 45 years old mean a lower absolute frequency by these phrases than do younger respondents. "Pretty often" is more often for whites than for blacks. Because the race difference for "very often" is borderline ($t = 1.87$, $p < .06$) and the number of blacks in this sample is small, however, future studies should examine the possibility of other race differences in meaning. None of the sex differences in the meaning of the relative-frequency categories is significant. (For sex, a model that includes only the category variables and constrains the meaning of "pretty often" and "very often" to be the same for excitement and boredom fits significantly better [$p < .01$] than the model in table 4. The estimates of the coefficients in the model are only .01 smaller than the corresponding estimates in table 4.)[10]

---

[9] For raw absolute frequencies, the pattern of significant results is almost identical. The following significance tests for boredom are different: There is a significant (t = 2.06) race difference in the frequency implied by "very often" and a significant (t = -2.42) sex difference in the meaning of "pretty often." For age, the difference in the frequency implied by "not too often" for those aged 31–45 and that in the frequency implied by "very often" for those aged 61 or older are not significant (t = $-.82$ and t = -1.61, respectively). There is no difference in the overall number of significant group differences. If an effect is significant when raw frequencies are analyzed but not when logs are analyzed, the group difference is probably greatest in the use of high values.

The prediction that relative frequencies should imply a greater absolute frequency for excitement than boredom is confirmed only for "not too often." The prediction that a given relative frequency should mean a greater absolute frequency among the group with the greater overall frequency of an event fares better.[11] This prediction generally describes the use of "pretty often" and "very often": for both education and age, for example, the group differences in the meaning of the relative frequencies (table 4) consistently mirror the pattern of overall group mean differences in the sample (table 1). The patterns are similar even when the overall group means do not differ significantly from each other.

These results can be used to understand why the conclusions suggested by absolute and relative frequencies sometimes differ. If relative frequencies reflect an implicit comparison to an expected frequency that varies by social group, using relative frequencies should

---

10. Analyzing the raw absolute frequencies also leads to accepting the model in table 4. The results of significance tests differ from those in table 4 in the following ways when raw frequencies are analyzed: For race, the difference in the meaning implied by "very often" is significant ($t = 2.26$, $p = .02$). For education, the difference in the frequency implied by "pretty often" for high school graduates is borderline ($t = 1.74$, $p = .08$). For age, the difference in the meaning implied by "not too often" for those aged 31–45 is not significant ($t = -.64$, $p = .52$), and the difference in the meaning implied by "pretty often" for those over 60 is borderline ($t = -1.87$, $p = .06$).

11. The group differences found here appear only at the higher relative frequencies, where the variation in meaning (the variance of the absolute frequencies) is greater. It may be that characteristics of the event dominate the use of low relative frequencies (where the variation in meaning is less).

Table 3. Goodness-of-Fit Statistics for Selected Constraints on the Basic Model

A. Constraints on the Model

		Group Variable							
		Race		Sex		Education		Age	
Model	Constrained Effects	$-2\log(L)$	k	$-2\log(L)$	k	$-2\log(L)$	k	$-2\log(L)$	k
A	G, C, G^*C	560	8	600	8	548	12	529	16
B	G, (PO, VO), G^*C	524	9	564	9	512	13	489	17
C	G, (PO, VO), (G^*PO, G^*VO)	524	10	564	10	507	15	483	20
D	G, (PO, VO)	521	12	564	12	499	19	478	26
E	G	520	14	563	14	497	21	474	28
F	Unconstrained	520	15	562	15	497	23	474	31

B. Selected Likelihood Ratio Tests

	Group Variable											
	Race			Sex			Education			Age		
Contrasts between Models	L^2	*df*	*p*	L^2	*df*	*p*	L^2	*df*	*p*	L^2	*df*	*p*
A vs. B	37	1	.00	36	1	.00	37	1	.00	40	1	.00
B vs. C	0	1	.99	0	1	.80	5	2	.10	5	3	.14
C vs. D	2	2	.32	0	2	.98	8	4	.08	5	6	.51
D vs. E	1	2	.53	1	2	.53	2	2	.34	4	2	.12
E vs. F	0	1	.99	0	1	.58	0	2	.95	0	3	.50

Note.—The constraints specify that the effect named is the same for the equations predicting excitement and boredom. In panel A, G refers to the set of group dummy variables, C to the set of relative frequency category dummies, and $C*G$ to the set of their interactions. All models constrain the intercept to be the same in both equations. The equality of two of the relative frequency categories (C)—pretty often (PO) and very often (VO)—is tested separately. Likelihood ratio statistics and probability levels were computed before rounding. k = number of parameters in the model. L = likelihood of the model. L^2 = likelihood ratio test statistic.

Table 4. Coefficients for Constrained Models Predicting the Absolute Frequency of Being Excited and Bored from Relative Frequency Category and Race, Sex, Education, and Age

	Race		Sex		Education		Age	
	b	*t*	b	*t*	b	*t*	b	*t*
Intercept (never)	−.015	−.37	−.001	−.03	−.006	−.26	.002	.08
Relative-frequency category:								
Excited not too often	.746	15.07	.742	27.81	.713	23.60	.779	22.58
Bored not too often	.623	13.09	.622	27.04	.594	21.50	.656	21.09
Pretty often	.922	17.98	1.033	41.11	.932	29.74	1.100	33.72
Very often	1.068	20.57	1.159	41.22	1.050	31.16	1.242	34.32
Group variable(s):								
G_1	.017	.40	−.002	−.08	.007	.24	−.009	−.26
G_2	. . .	. . .	. . .	. . .	.009	.28	−.016	−.43
G_3	. . .	. . .	. . .	. . .	. . .	. . .	−.002	−.05

Category by group interactions:								
Not too often * G_1	−.013	−.26	−.019	−.61	.038	1.03	−.074	−1.79
Not too often * G_2	. . .	. . .	. . .	. . .	.012	.31	−.030	−.68
Not too often * G_3	. . .	. . .	. . .	. . .	. . .	. . .	−.080	−1.81
Pretty often * G_1	.110	2.02	−.027	−.81	.077	1.83	−.088	−1.96
Pretty often * G_2	. . .	. . .	. . .	. . .	.154	3.68	−.165	−3.48
Pretty often * G_3	. . .	. . .	. . .	. . .	. . .	. . .	−.133	−2.59
Very often * G_1	.105	1.87	−.008	−.21	.126	2.70	−.038	−.74
Very often * G_2	. . .	. . .	. . .	. . .	.188	3.96	−.184	−3.40
Very often * G_3	. . .	. . .	. . .	. . .	. . .	. . .	−.231	−4.00
Rho	.30		.30		.30		.29	

NOTE.—For race, the reference category is black and G_1 is white. For education, the reference category is less than high school, G_1 is high school graduate, and G_2 is more than high school education. For age, the reference category is 30 or younger, G_1 is 31–45, G_2 is 46–60, and G_3 is 61 or older. The dependent variable is the log of absolute frequencies, with 1 added before taking logs. The intercept and effects for "pretty often" and "very often," for the group variables, and the group by category interactions are constrained to be the same for excitement and boredom.

reduce group differences. (This would be true, of course, only if other social and response processes generating relative frequency reports were similar for the groups.) Because the relative and absolute metrics are inherently incomparable, however, it is difficult to evaluate whether this is generally true in table 1. Comparing the relative and absolute frequency of excitement for blacks and whites does suggest that relative frequencies reduce the group difference: whites mean a higher log absolute frequency by "pretty often" and (in the sample) "very often," and so the distribution of whites and blacks across relative frequency categories is very similar for excitement.

In one case the group for which the relative frequencies imply a higher absolute frequency is not the group with the higher overall mean absolute frequency. This occurs for the comparison of race differences in the frequency of boredom. The overall mean absolute frequency is quite similar for whites and blacks. But whites persist in meaning a greater absolute frequency by "pretty often," and (in the sample) "very often," and the relative frequency of boredom is significantly greater among blacks than whites. This could be explained as a question order effect: whatever meaning the respondent assigns the relative frequencies when reporting about excitement may influence the subsequent report for boredom. Such a carryover would be masked for education and age, because the group means for the absolute frequencies are ordered in the same way for excitement and boredom. Alternatively, factors that cannot be examined here may affect the meaning of relative frequency phrases. Such factors probably contribute to the variation of the absolute frequencies within each of the relative frequency categories. This variation can also cause the results obtained using relative and absolute frequencies to differ.

In conjunction with earlier research, finding that the meaning of some relative frequencies differs by race, education, and age suggests that constructing relative frequencies may involve simultaneous comparisons with several social groups (in addition to other evaluations not examined here). Young well-educated blacks, for example, may implicitly compare themselves to others like themselves in deciding whether the number of times they were bored is "pretty often."[12]

---

[12] Although no significant difference in meaning by sex was found here, sex differences should be examined for other events where the distinction between relative and absolute frequency is potentially theoretically important. Compare, e.g., the response scales used by the Psychiatric Evaluation Research Interview and the Center for Epidemiological Studies' depression scale (Newmann 1984; Radloff 1977; Ross and Mirowsky 1984). Even differences in meaning as small as those observed here for males and females can cumulate when items are combined in a scale. For example, if all items in the positive affect scale included in this study are scored using pooled means of the absolute frequencies or equal ad hoc intervals, females appear to show significantly greater positive affect than males. If the items are scored using sex-specific means, the difference reverses sign

Discussion

Differences in reports of absolute and relative frequencies can be conceptualized in several ways. Perhaps individuals do not store accurate information about the absolute frequency of events, and relative frequencies are the best frequency information respondents have to offer (Bradburn and Danis 1984, p. 125). But to a researcher hoping that relative frequencies are a useful alternative to absolute frequencies, the variability in their meaning might be evidence that relative frequencies entail more measurement error than absolute frequencies (Groves 1989, pp. 452–54). The literature reviewed in the introduction suggests a compromise. The two types of frequencies may be substantively different and each subject to characteristic errors.

Absolute frequencies appear to be inherently "harder" data than relative frequencies—counts of events or behaviors are less ambiguous, are potentially verifiable, and have comparable metrics in different groups. The known systematic errors in self-reported absolute frequencies may, however, compromise their potential advantages (and make them a risky criterion to use in evaluating relative frequencies). Experimental subjects appear to overestimate low frequencies and underestimate high frequencies for both experimentally controlled low-frequency events and naturally occurring high-frequency events (Bradburn and Danis 1984). Other factors that affect reports of absolute frequency include the ease with which instances of the event are remembered (Bradburn, Rips, and Shevell 1987; Sherman and Corty 1984; Tversky and Kahneman 1977); how often the person thinks about or imagines instances (Johnson, Taylor, and Raye 1977); how well the person imagines instances (Johnson et al. 1979); the spacing of events and the stability or variation in the context in which they occur (Rose 1980); recall techniques that the respondent uses (Blair and Burton 1987); feelings at the time of the event and at the time of recall (Isen 1984); response categories (Schwarz and Hippler 1987); and the true frequency (e.g., Belson 1986, p. 134).

Another kind of error may be present in some respondents' reports of the absolute frequency of excitement and boredom: such internal events pose a very difficult response task because the events have vague boundaries and are difficult to remember. To simplify the task, some respondents may consult an intensity dimension. In the present study, language encourages confounding intensity and frequency.

---

and is no longer significant. (See Bradburn and Miles [1979] for examples of the scoring method and Schaeffer [1988].) If the focus of an analysis is absolute frequency and the meaning of relative frequencies differs by group, additive scales should not score categories the same way for all groups.

"Bored," like "excited," can refer either to an internal event (e.g., being bored) or to an intensity dimension of the same name (e.g., how bored a person felt). In addition, both the frequency of being bored and the intensity of the boredom probably play a role in judgments of the level of "boredom" present during a given span of time. (For an illustrative argument on a related topic, see Diener, Sandvik, and Pavot [1991].) Absolute frequencies can easily be used as "magnitude estimates" to express an intensity response (e.g., Lodge 1981). Depending on the kind of event asked about, some of those who produce a "direct estimate" of an absolute frequency may be adapting their answers in this way (Blair and Burton 1987).

One might argue that, despite their problems, absolute frequencies avoid the ambiguities of relative frequencies. This argument is strongest if one assumes that respondents store accurate absolute frequencies and introduce translation errors when reporting relative frequencies. Some of the ambiguity of relative frequencies may be unavoidable, however, if (or when) the translation proceeds in the other direction. That is, respondents may use stored relative-frequency information along with heuristics to compute absolute frequencies. Decomposition strategies, which may use relative frequencies to identify the categories of the decomposition, and respondents' aversion for episodic enumeration are compatible with an initial retrieval of relative frequencies (Blair and Burton 1987; Bradburn, Rips, and Shevell 1987; Schwarz 1990). So is the slightly higher rate of missing information on the questions asking for absolute frequencies: of those who provided a relative frequency for excitement, 7 percent gave no absolute frequency; the corresponding percentage for the question about boredom is 5 percent.

Previous research does not conclusively describe the relationship between absolute and relative frequencies. Whether researchers prefer relative or absolute frequencies may depend on how they view the disadvantages of each and their relationship to each other. Research on errors in reports of absolute frequencies indicates that such reports are not simple transcriptions of some objective reality. But requests for absolute frequencies suggest that respondents count, estimate, or average the actual number of behaviors or events, and most people at least attempt to do so (Blair and Burton 1987). In contrast, a request for relative frequencies implicitly invites a larger evaluative component: respondents are free to choose their own standards (e.g., in deciding what is "often"), and the category labels are similar to those used to obtain subjective evaluations along intensity dimensions.[13] The research cited in the introduction and some of the findings reported here support viewing relative frequencies as at least partly the result of evaluations for which expected frequency and feelings about the event are two relevant inputs.

---

[13] The conventions of language use may be one source of error in relative frequencies as surrogates for absolute frequencies. Relative frequencies are routinely used rhetorically in "extreme case formulations" such as "I *always* pay cash" (Pomerantz 1986). If a researcher asks about an event that a respondent commonly formulates in this

The variability in the meaning of relative frequencies (as reflected in the variation of absolute frequencies within relative-frequency category) suggests that relative frequencies are poor substitutes for absolute frequencies. This is true whether the variation arises (1) because different respondents assign different relative frequencies to the reasonably accurate absolute frequencies they retrieve, (2) because the absolute frequencies respondents estimate for the relative frequencies they retrieve are subject to considerable systematic and random error, or (3) both. For both relative and absolute measures, errors probably increase as frequency increases, because recall is more difficult and because the size of errors for low frequencies is artificially constrained. The research cited earlier, for example, suggests that the absence of an event may be reported more accurately than other frequencies (Belson 1986) and that there is fairly high consensus on the meaning of "never" (Hakel 1969; Simpson 1944).[14] Once we understand the processes generating relative frequencies, we may find that relative and absolute frequencies are each more useful for some purposes than others.

Conclusion

In some respects the study of social phenomena using self-reports obtained by standardized questions is the study of how respondents use the specialized vocabulary presented to them in surveys. The analyses presented here suffer from the problems of secondary analysis—the data set was not designed to address the questions posed of it here. The results suggest that the meaning of relative frequencies can vary by group and that the choice of absolute or relative frequencies may affect comparisons. Which group comparisons are affected may depend on the event. For some events, group membership probably identifies differences in the frequency of events or specific norms that guide their evaluation or expression in language. Using a very different strategy, for example, Smith, Furstenberg, and Hughes (1990) find that the meaning of "just a few, if any" differs for blacks and whites.

---

way, the respondent may produce the rhetorical answer rather than an actual frequency estimate.

[14] Dichotomizing relative frequencies as "never" versus "ever" may thus sometimes be appropriate for group comparisons when even limited information about absolute frequencies is preferred to information about relative frequencies.

Moreover, we cannot assume that the intensifiers used in questions about subjective phenomena are immune to such group differences in meaning. We need to understand social sources of meaning before we can assess how group differences in meaning might affect research conclusions.

Research by cognitive psychologists suggests that absolute and relative responses do not simply represent objective and subjective realities; they are differently constructed perceptions that probably have different error structures. The circumstances under which absolute or relative responses are preferable cannot be specified in any simple way. When group comparisons are planned, the measurement theory that guides question design must be augmented with theories about how respondents make absolute and relative judgments.

Appendix

Table A1. Summary of Distributions of the Absolute Frequency of Being Excited and Being Bored, by Relative-Frequency Category

	Excited			Bored		
	Not Too Often	Pretty Often	Very Often	Not Too Often	Pretty Often	Very Often
Minimum	1	1	1	1	1	1
Deciles:						
1	1	2	4	1	3	3
2	2	4	5	1	4	5
3	2	5	8	2	4	8
4	4	8	12	2	8	12
5	4	8	16	3	12	16
6	4	12	20	4	12	20
7	8	16	24	4	20	24
8	8	20	28	6	28	28
9	15	28	30	8	28	30
Four largest values[a]	32 (1)	60 (2)	56 (2)	28 (2)	31 (1)	28 (14)
	46 (1)	80 (2)	80 (3)	30 (1)	32 (1)	30 (10)
	56 (1)	90 (1)	95 (1)	40 (1)	40 (2)	56 (1)
	97 (1)	97 (1)	97 (2)	97 (1)	48 (1)	96 (1)
Number of different values	22	28	24	18	20	23

Table A1. (Continued)

	Excited			Bored		
	Not Too Often	Pretty Often	Very Often	Not Too Often	Pretty Often	Very Often
Four largest modes[b]	1 (14)	4 (15)	4 (12)	1 (22)	4 (19)	12 (13)
	2 (19)	8 (19)	8 (13)	2 (28)	8 (14)	20 (14)
	4 (21)	12 (12)	12 (13)	4 (23)	12 (12)	28 (14)
	8 (15)	16 (10)	28 (17)	8 (12)	28 (16)	30 (10)
Raw frequencies:						
Mean	6.53	12.94	17.53	4.05	13.72	16.99
S.D.	7.48	12.02	14.42	4.39	10.64	11.06
Skewness	3.19	2.49	2.06	3.92	.77	.74
Log frequencies:						
Geometric mean	4.17	8.91	12.30	2.88	9.33	12.59
Mean	.62	.95	1.09	.46	.97	1.10
S.D.	.40	.40	.41	.35	.42	.40
Skewness	.25	−.22	−.71	.46	−.50	−1.01
N	327	495	247	552	127	99

NOTE.—Values of 97 or greater are coded 97 in the original data. For the raw frequencies, seven values greater than 90 were recoded to the next lowest observed value within the relative-frequency category before calculating means, standard deviations, and skew. This transformation was also used for all statistics and analyses reported in the text. In addition, when statistics or analyses in the text include those with a relative frequency of "never," the absolute frequency of "never" is set to zero, and one is added before taking logs.

[a] Number of respondents choosing the value is in parentheses.

[b] Numbers in parentheses are percentages of respondents choosing the value.

References

Bass, Bernard M. 1968. "How to Succeed in Business according to Business Students and Managers." *Journal of Applied Psychology* 52:254–62.

Bass, Bernard M., Wayne F. Cascio, and Edward J. O'Connor. 1974. "Magnitude Estimation of Expressions of Frequency and Amount." *Journal of Applied Psychology* 59:313–20.

Belson, William A. 1986. *Validity in Survey Research*. Aldershot: Gower.

Blair, Edward, and Scot Burton. 1987. "Cognitive Processes Used by Survey Respondents to Answer Behavioral Frequency Questions." *Journal of Consumer Research* 14:280–88.

Bradburn, Norman M. 1969. *The Structure of Psychological Well-Being*. Chicago: Aldine.

Bradburn, Norman M., and Catalina Danis. 1984. "Potential Contributions of Cognitive Research to Survey Questionnaire Design." In *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*, ed. Thomas Jabine, Miron Straf, Judith Tanur, and Roger Tourangeau, pp. 101–29. Washington, DC: National Academy Press.

Bradburn, Norman M., and Carrie Miles. 1979. "Vague Quantifiers." *Public Opinion Quarterly* 43:92–101.

Bradburn, Norman, Lance J. Rips, and Steven K. Shevell. 1987. "Answering Autobiographical Questions: The Impact of Memory and Inference on Surveys." *Science* 236:157–61.

Bradburn, Norman M., Seymour Sudman, and Associates. 1979. *Improving Interview Method and Questionnaire Design*. San Francisco: Jossey-Bass.

Chase, Clinton I. 1969. "Often Is Where You Find It." *American Psychologist* 24:1043.

Cliff, Norman. 1959. "Adverbs as Multipliers." *Psychological Review* 66:27–44.

Dennis, Russell, Cathy Dennis, and Larry Jobe. 1975. "The Meaning of Frequency Words in Different Contexts." *Language and Speech* 18:343–49.

Diener, Ed, Ed Sandvik, and William Pavot. 1991. "Happiness Is the Frequency, Not the Intensity, of Positive versus Negative Affect." In *Subjective Well-Being: An Interdisciplinary Approach*, ed. Fritz Strack, Michael Argyle, and Norbert Schwarz, pp. 119–39. Oxford: Pergamon.

Felmlee, Diane H., and Lowell L. Hargens. 1988. "Estimation and Hypothesis Testing for Seemingly Unrelated Regressions: A Sociological Application." *Social Science Research* 17(4):384–99.

Greene, William. 1988. LIMDEP Version 5.

Goocher, Buell E. 1965. "Effects of Attitude and Experience on the Selection of Frequency Adverbs." *Journal of Verbal Learning and Verbal Behavior* 4:193–95.

Groves, Robert M. 1989. *Survey Errors and Survey Costs*. New York: Wiley.

Hakel, Milton. 1969. "How Often Is Often?" *American Psychologist* 23:533–34.

Isen, Alice M. 1984. "Toward Understanding the Role of Affect in Cognition." In *Handbook of Social Cognition*, ed. Robert Wyer and Thomas Srull, 3:179–236. Hillsdale, NJ: Erlbaum.

Johnson, Marcia K., Carol L. Raye, Alvin Y. Wang, and Thomas H. Taylor. 1979. "Fact and Fantasy: The Roles of Accuracy and Variability in Confusing Imaginations with Perceptual Experiences." *Journal of Experimental Psychology: Human Learning and Memory* 5:229–40.

Johnson, Marcia K., Thomas H. Taylor, and Carol L. Raye. 1977. "Fact and Fantasy: The Effects of Internally Generated Events on the Apparent Frequency of Externally Generated Events." *Memory and Cognition* 5:116–22.

Jones, Lyle V., and L. L. Thurstone. 1955. "The Psychophysics of Semantics: An Experimental Investigation." *Journal of Applied Psychology* 39:31–36.

Lichtenstein, Sarah, and Robert Newman. 1967. "Empirical Scaling of Common Verbal Phrases Associated with Numerical Probabilities." *Psychonomic Science* 9(10):562–64.

Lodge, Milton. 1981. *Magnitude Scaling: Quantitative Measurement of Opinions*.

Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-025. Beverly Hills, CA: Sage.

Lodge, Milton, David Cross, Bernard Tursky, Joseph Tannenhaus, and Richard Reeder. 1976. "The Psychophysical Scaling of Political Support in the 'Real World.'" *Political Methodology* 3:159–82.

Maddala, G. S. 1977. *Econometrics*. New York: McGraw-Hill.

Mosier, Charles I. 1941. "A Psychometric Study of Meaning." *Journal of Social Psychology* 13:123–40.

Newmann, Joy P. 1984. "Sex Differences in Symptoms of Depression." *Journal of Health and Social Behavior* 25:136–59.

Newstead, Stephen E., and Janet M. Collis. 1987. "Context and the Interpretation of Quantifiers of Frequency." *Ergonomics* 30(10):1447–62.

Pace, C. Robert, and Jack Friedlander. 1982. "The Meaning of Response Categories: How Often Is 'Occasionally,' 'Often,' and 'Very Often'?" *Research in Higher Education* 17(3):267–81.

Parducci, Allen. 1982. "Category Ratings: Still More Contextual Effects!" In *Social Attitudes and Psychophysical Measurement*, ed. Bernd Wegener, pp. 89–105. Hillsdale, NJ: Erlbaum.

Pepper, Susan. 1981. "Problems in the Quantification of Frequency Expressions." In *New Directions for Methodology of Social and Behavioral Sciences*, no. 9, ed. Donald Fiske, pp. 25–41. San Francisco: Jossey-Bass.

Pepper, Susan, and Lubomir S. Prytulak. 1974. "Sometimes Frequently Means Seldom: Context Effects in the Interpretation of Quantitative Expressions." *Journal of Research in Personality* 8:95–101.

Pohl, Norval F. 1981. "Scale Considerations in Using Vague Quantifiers." *Journal of Experimental Education* 49(4):235–40.

Pomerantz, Anita. 1986. "Extreme Case Formulations: A Way of Legitimizing Claims." *Human Studies* 9:219–29.

Radloff, Lenore. 1977. "The CES-D Scale: A Self-report Depression Scale for Research in the General Population." *Applied Psychological Measurement* 1(3):385–401.

Response Effects in Surveys: 1975 National Study (machine-readable data file). 1980. Norman Bradburn and Seymour Sudman, principal investigators. Chicago: National Opinion Research Center. (One data file [1,172 logical records] and SPSS data definition commands, plus accompanying documentation.)

Rose, Robert J. 1980. "Encoding Variability, Levels of Processing, and the Effects of Spacing of Repetitions upon Judgments of Frequency." *Memory and Cognition* 8:84–93.

Ross, Catherine E., and John Mirowsky. 1984. "Components of Depressed Mood in Married Men and Women." *American Journal of Epidemiology* 119:997–1004.

Rubin, Donald B. 1984. "Distinguishing between the Scale of the Estimand and the Transformation to Normality." *Journal of the American Statistical Association* 79:309–12.

Schaeffer, Nora Cate. 1988. "Sex Differences in the Use of Vague Quantifiers." Paper presented at the annual meeting of the American Association for Public Opinion Research, Toronto, May.

———. 1989. "The Frequency and Intensity of Parental Conflict: Choosing Response Dimensions." *Journal of Marriage and the Family* 51(3):759–66.

Schaeffer, Nora Cate, and Norman M. Bradburn. 1989. "Respondent Behavior in Magnitude Estimation." *Journal of the American Statistical Association* 84(406):402–13.

Schriesheim, Chester A., and Janet F. Schriesheim. 1978. "The Invariance of Anchor Points Obtained by Magnitude Estimation and Pair-Comparison Treatment of Complete Ranks Scaling Procedures: An Empirical Comparison and Implications for Validity of Measurement." *Educational and Psychological Measurement* 38:977–83.

Schwarz, Norbert. 1990. "Assessing Frequency Reports of Mundane Behaviors: Contributions of Cognitive Psychology to Questionnaire Construction." In

Research Methods in Personality and Social Psychology, ed. Clyde Hendrick and Margaret S. Clark. Review of Personality and Social Psychology, vol. 11. Beverly Hills, CA: Sage.

Schwarz, Norbert, and Hans-J. Hippler. 1987. "What Response Scales May Tell Your Respondents: Informative Functions of Response Alternatives." In *Social Information Processing and Survey Methodology*, ed. Hans-J. Hippler, Norbert Schwarz, and Seymour Sudman, pp. 163–78. New York: Springer-Verlag.

Sherman, Steven J., and Eric Corty. 1984. "Cognitive Heuristics." In *Handbook of Social Cognition*, ed. Robert S. Wyer and Thomas K. Srull, 1:189–286. Hillsdale, NJ: Erlbaum.

Simpson, Ray H. 1944. "The Specific Meanings of Certain Terms Indicating Differing Degrees of Frequency." *Quarterly Journal of Speech* 30:328–30.

Smith, Herbert L., Frank F. Furstenberg, and M. E. Hughes. 1990. "Variation in Mother-Teenager Agreement in a Survey of Philadelphia Area Parents and Teens." Unpublished manuscript, Population Studies Center, University of Pennsylvania.

Spector, Paul E. 1976. "Choosing Response Categories for Summated Rating Scales." *Journal of Applied Psychology* 61:374–75.

Torgerson, Warren S. 1958. *Theory and Methods of Scaling*. New York: Wiley.

Tversky, Amos, and Daniel Kahneman. 1977. "Availability: A Heuristic for Judging Frequency and Probability." *Cognitive Psychology* 5:207–32.

Wallsten, Thomas S., David V. Budescu, Amnon Rapoport, Rami Zwick, and Barbara Forsythe. 1986. "Measuring the Vague Meaning of Probability Terms." *Journal of Experimental Psychology: General* 115:348–65.

THE BUILDING OF A REPUBLICAN REGIONAL BASE IN THE SOUTH

THE ELEPHANT CROSSES THE MASON-DIXON LINE

MARTIN P. WATTENBERG

Abstract Changes in party identification are examined from 1952 to 1988, with a special emphasis on the findings from the 1988 CPS Super Tuesday Study. It is found that in 1988 Republicans clearly outnumbered Democrats among southern whites for the first time ever. Demographic patterns of partisanship in 1988 are compared to those from the 1950s. Most important is the relationship which now exists between age and party identification. Republicans now have a 29 percent lead over the Democrats among southern whites under 30 years of age. While firm black support of the Democratic party should prevent the South from becoming as solidly Republican as it once was Democratic, there is good reason to expect further Republican gains in the South throughout the remainder of the twentieth century.

The 1988 election of George Bush marked the second time in a row that the South voted 4 percent above the national average for the Republican presidential ticket. Just as the Democrats started every presidential campaign during the New Deal with the head start of having the South locked up, it now appears that the Republicans enjoy the strategic benefits of having a strong regional base. If the Republicans can count on winning most of the South's electoral votes, as well as those of the Mountain West (see Galderisi et al. 1987), they need only concentrate on winning a few of the largest states to win a presidential election. Documenting how the Democratic South became the Republican South is therefore crucial to understanding how the Republicans, once accustomed to losing presidential elections, have now become the steady favorite.

MARTIN P. WATTENBERG is professor of political science at the University of California, Irvine. The data presented in this article are from the American National Election Studies, 1952–88, and were provided by the Inter-University Consortium for Political and Social Research.

Public Opinion Quarterly Volume 55:424–431 © 1991 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/91/5503-0009/$2.50

Party Identification in the South: 1952–88

Realignment in the South has been a process of secular rather than sudden change. While election returns took a decided change in the Republican direction starting in 1964 (see Cosman 1966), it was not until 1988 that the Republicans clearly had the edge in party identification among southern whites. Table 1 displays the changes from 1952 to 1988 by race for the two regions known in the National Election Studies as the Solid and Border South.[1] The two regions are combined in this analysis due to the consistent finding that many people in the Border states consider themselves southerners.[2] For example, 51 percent of those interviewed in the Border South in 1984 said southerners were a group they felt close to in terms of their ideas and feelings about things, compared to 46 percent in the Solid South and just 17 percent in the non-South.

A second definitional choice which must be noted in table 1 is the incorporation of independent Democrats and independent Republicans along with the partisan identifiers. According to Keith et al. (1977) independent leaners have displayed voting patterns which are virtually identical to weak partisans. Indeed, independent Republicans in the South voted an average of 93.6 percent Republican from 1972 to 1984 compared to 90.6 percent for weak Republicans. To not take into account the leaners would thus be to underestimate the rise of southern Republicanism.

In his analysis of party identification in the South from 1952 to 1960, Converse (1966) noted that, should the trends in the South and non-South continue at the same rate, the point of convergence would occur in June 1983. While Converse's prediction was figurative rather than literal, it turned out to be quite accurate. The rate of change slowed considerably in the South after the 1964 election, but the onset of the Reagan presidency led to the fulfillment of the prophecy.

During Reagan's first term in office the percentage of Republicans among white southerners increased from 33.3 percent in 1980 to 45.9 percent in 1984, as can be seen in table 1.[3] The 1984 election year

[1] The American National Election Studies sample of the Solid South includes Virginia, Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, and Texas; the Border South is defined as Kentucky, Maryland, Oklahoma, Tennessee, Washington, D.C., and West Virginia.

[2] In this section of the American National Election Study respondents are given a list of groups and are asked to tell which groups they "feel particularly close to—people who are most like you in their ideas and interests and feelings about things."

[3] The 1984 continuous monitoring study is used here rather than the traditional pre-post study for two reasons. First, the former has a much larger number of cases in the region (1,793 vs. 725) and therefore a smaller sampling error. The second is that there was an

Table 1. Party Identification in the South: 1952–88 (Independent Leaners Are Included as Partisans)

	1952	1956	1960	1964	1968	1972	1976	1980	1984	1988
Whites:										
Democrats	75.8	67.1	64.5	68.3	57.9	54.1	54.0	52.5	44.4	40.6
Independents	2.6	5.6	10.8	6.6	13.3	12.4	15.1	12.3	9.1	5.6
Republicans	18.3	23.6	22.9	24.0	27.3	31.5	30.4	33.3	45.9	51.1
Apoliticals	3.4	3.7	1.8	1.1	1.5	1.9	.5	1.9	.5	2.7
Blacks:										
Democrats	57.0	48.2	41.5	82.1	93.6	73.2	82.7	82.3	78.6	85.0
Independents	.9	2.4	11.0	7.4	1.3	14.0	10.2	5.2	6.9	4.7
Republicans	18.4	18.1	23.2	4.2	1.3	11.5	6.6	8.3	13.7	8.3
Apoliticals	23.7	31.3	24.4	6.3	3.8	1.3	.5	4.2	.8	2.1

SOURCES.—1952–80 SRC/CPS National Election Studies; 1984 CPS Continuous Monitoring Study; 1988 CPS Super Tuesday Study.

marked the first time that Republicans outnumbered Democrats among white southerners, though the margin was well within sampling error. By Super Tuesday of 1988 the Republican's dominance was clear, however. In the 13 out of the 15 southern states in which interviews were conducted prior to Super Tuesday, 51.1 percent of white southerners identified themselves as Republicans compared to just 40.6 percent as Democrats. While this study does not constitute a perfect sample of the South in that South Carolina and West Virginia were excluded, the 13 states which were sampled represent 94 percent of the region's population. An analysis of survey data from 1980 to 1986 combined shows that when South Carolina and West Virginia are taken out of the analysis, the percentage of Republicans decreases by a mere 0.5 percent. Thus, it can safely be said that the 1988 CPS Super Tuesday survey is representative of the South as a whole. What it suffers from in terms of sampling frame is more than made up for by the large number of 1,838 southern respondents. Because of this large sample size, the 1988 Super Tuesday study offers an unparalleled opportunity to examine the correlates of southern partisanship at a major turning point in the region's political history.

---

apparent bias in the pre-post 1984 study toward the Democrats. Whereas the election returns from the South show that 62 percent of the votes cast were for Reagan, the post-election sample reported only a 54 percent Reagan vote. Thus, there clearly was some serious sampling problem in the South in this study—one which seems to have been repeated in the 1988 pre-post study as well.

Table 2. Party Identification of White Southerners by Demographic Groups (Entries Are % Democrats Minus % Republicans, including Leaners)

	1988	1952–56		1988	1952–56
Section:			Sex:		
Solid South	−12.6	+60.4	Male	−18.2	+59.3
Border South	−3.7	+23.5	Female	−3.9	+45.9
Age:			Income:[a]		
Under 30	−28.0	+45.7	Low	+14.4	+39.0
30–44	−10.1	+57.6	Lower middle	−12.6	+51.9
45–64	−2.8	+49.5	Middle	−16.3	+63.5
65 and over	−.4	+52.6	Upper middle	−26.7	+56.8
Religion:			High	−8.0	+38.8
Protestant	−14.0	+52.5	Union household:		
Catholic	−2.9	+47.2	Union	+19.5	+56.7
Jewish	+29.0	. . .[b]	Nonunion	−13.9	+51.0
			Education:		
			Grade school	+20.5	+43.1
			High school	+1.1	+57.5
			Some college	−14.8	+55.8
			College graduate	−31.2	+54.1

SOURCES.—1952–56 SRC National Election Studies; 1988 CPS Super Tuesday Study.

[a] Income categories by percentiles are approximately as follows: 0–25, 26–45, 46–66, 67–96, and 97–100.

[b] Only nine Jews were interviewed in the South in this period.

Changes in White Southern Partisanship by Demographic Groups

In the 1950s the simple fact of being a southerner was sufficient reason for most to identify as Democrats; many had been born into the Democratic party in much the same way they had been born into a given church. Their political identity was overwhelmingly defined by where they lived, as opposed to who they were in terms of religion, social status, etc. This absence of major demographic patterns in the partisanship of whites in the old South is shown in table 2 by pooling the data from 1952 and 1956. In contrast, comparable data from the 1988 Super Tuesday study show many significant relationships in the new South, which now virtually match those in the rest of the country.

The most important 1988 pattern in table 2 is clearly that of age. The fact that citizens under 30 are by far the most Republican whereas those over 65 are the most Democratic should benefit the GOP for a generation to come, as population replacement takes its toll.[4] Yet such a pattern is not unique to the South. As Norpoth (1987) has shown, in the 1980s young people have come to identify with the Republicans more than their elders. His analysis outlines the contrast to previous Republican victories, in which young voters were the most Democratic and older voters—socialized before the New Deal—were the most Republican. In the old South, however, such an age pattern never existed, as one would have had to been socialized before the Civil War to remember a time when the Democrats were not dominant.

Another aspect of the New Deal party system which made little imprint on the old South was the role of social status, as measured by either family income or education. As Converse (1966, p. 221) wrote, the Democrats in the old South were "the symbol of small-town, middle-class respectability." Thus table 2 shows that in the 1950s low income and a grade school education actually decreased the likelihood of Democratic identification among white southerners in the 1950s. Such an anomaly helps to explain Key's (1949) findings regarding the effect of the southern one-party system on redistributive policy. Given that the poor and grade school educated were less likely to vote, as well as less suportive of the dominant party, it is hardly surprising that they received few benefits from the system.

Key predicted that as the South industrialized socioeconomic patterns of partisanship would come to resemble those of the North. As can be seen in table 2, the 16 percent of white southerners who belonged to a union household in the 1950s were slightly more likely to be Democrats. By 1988 this unionized group had declined to 9 percent, but the difference in partisanship compared to the rest of the population had increased greatly. Thus, what little remaining strength unions have in the South has contributed to holding it from becoming even more Republican. In fact, people living in union households are one of the few groups of white southerners among which the Democrats still have an edge.

As the South has moved beyond industrialization to high technology, increasing education and income levels have substantially contributed to the GOP gains. The middle and upper-middle income categories are now major sources of Republican strength. At first glance, it seems peculiar that those in the top 4 percent of family income are somewhat less Republican than many with lesser incomes. This apparent enigma can be explained, however, by the rather advanced age of the wealthiest southern whites. Higher education, on the other hand, is more likely to be found among the young, and thus a more linear relationship between education and party can now be found than with income. As grade school–educated citizens have dwindled, this aging remnant of the old South has gone from the least to the most Democratic of the education categories. In contrast, among college graduates a 54 percent advantage for the Democrats in the 1950s has been transformed to a 31 percent edge for the Republicans in 1988. As Black and Black (1987, p. 247) have written, "'Going Republican' has become the central tendency of educated whites, and the tremendous increase in sympathy and support for Republicanism among the region's most highly educated citizens provides firm support for competitive, two-party politics." In sum, the role of socioeconomic status in the South has at last fallen in line with the long-established pattern of the New Deal coalition.

---

[4] The pattern of young people being the most likely to identify as Republicans is also shown in most of the Southern states analyzed individually in Swansbrough and Brodsky (1988).

Another demographic factor on which southern partisanship has become more like that of the rest of the nation is religious denomination. Outside the South an important cleavage has long existed between Protestant Republicans on the one side and Catholic and Jewish Democrats on the other. Yet, in the old South, Catholics constituted only 5 percent of the population and Jews such a negligible proportion that there were not enough cases for analysis. By 1988, though, Catholics had grown to 18 percent and Jews to 2 percent of the southern white population. As in the rest of the country, Catholics are now somewhat less supportive of the Republicans than are Protestants among whites, and Jews remain a rather pro-Democratic group.

In recent years the study of religion and politics has shifted in focus from denomination to evangelicalism. In particular, the new Christian right has often been identified as a southern-dominated movement (Patel, Pilart, and Rose 1982). Overall, the 1988 data show no partisan difference between white southern Protestants who consider themselves "born again" from those who do not. Controlling for age reveals an interesting pattern, however. Evangelical Christians are 8 percent more likely than other white Protestants to be Republicans among those under 30 and are 5 percent more Republican among those in the 30–44 age group. In contrast, the relationship is reversed among white southerners socialized prior to the civil rights revolution. Those who report being "born again" are 2 percent more likely to be Democrats in the 45–64-year-old-age range, and 7 percent more so in the over-65 group. This difference by age demonstrates a profound change in the political culture of the South—one in which evangelical preaching has changed from espousing Democratic to Republican views (see Guth 1985).

The effect of such religious beliefs is influenced by gender as well as age. Women are not only more likely to consider themselves "born again" but also are apparently more politically swayed by such beliefs. Except for the over-65 category, the impact on partisanship of being "born again" is greater for women than for men.

In spite of the greater incidence of deep religious belief among women, in all but the oldest age cohort they are now more Democratic than men in the white south. Such findings closely follow the nationwide gender gap of the 1980s. In the 1950s, the gender gap was reversed, with women being more Democratic in the North as well as the South. Thus, this is one aspect of partisanship in which the national pattern is clearly visible in the old as well as the new South.

Not only is the nation now more homogeneous in terms of partisanship, but so is the South. Whites in the Solid South have changed the most, moving from being more Democratic to more Republican than those in the Border South. The difference in partisanship between the two sections, however, has been cut to just a quarter of what it was in the 1950s. It thus makes more sense to combine the two than ever before.

A Solid Republican South?

Is any end in sight to the rise of Republicanism in the South? The age-related pattern of partisanship which currently exists among white southerners would point to the current trend having many more years to run. It is unlikely, though, that the South will ever become as solidly Republican as it once was Democratic. In the old South there was little social structuring of partisanship, which meant that the Republicans had no base of support upon which they could rely. This is clearly no longer the case. Most notably, southern blacks now provide the Democrats with a sizable base. To be black now defines a voter as a Democrat in much the same way that being a southerner defined one as a Democrat in the 1950s—little else about the individual matters in determining partisanship. With united black support, the Democrats need only between 30 and 40 percent of the white vote to achieve victory, which is within reach even among the youngest voters.

This is by no means to say that the Republican edge among young southern whites is of little consequence. Most important, they are not only calling themselves Republicans in survey interviews but registering as Republicans as well. Whereas registered Democrats continue to outnumber registered Republicans in every age group over 30, the Republicans now hold a registration advantage of 28–48 percent among southern whites under 30. This offers the Republicans a far larger pool of young partisans from which to draw their leaders of the future, especially given the current relationship between party identification and education. In particular, 80 percent of male college graduates under 30 in the white South have signed up as Republicans. Only 2 of the 41 respondents in this group reported registering as Democrats. And on party identification none called themselves "Strong Democrats," whereas 15 said they were "Strong Republicans." To the extent that well-educated white males continue to have an advantage in mounting well-financed campaigns, as well as in attaining the requisite experience to run for public office, the Republicans should be able to make significant gains at every level of southern politics throughout the remainder of the twentieth century.

References

Black, Earl, and Merle Black. 1987. *Politics and Society in the South*. Cambridge, MA: Harvard University Press.

Converse, Philip E. 1966. "On the Possibility of Major Political Realignment in the South." In Angus Campbell et al., *Elections and the Political Order*. New York: Wiley.

Cosman, Bernard. 1966. *Five States for Goldwater: Continuity and Change in Southern Presidential Voting Patterns*. University: University of Alabama Press.

Galderisi, Peter F., Michael S. Lyons, Randy T. Simmons, and John G. Francis. 1987. *The Politics of Realignment: Party Change in the Mountain West*. Boulder, CO: Westview.

Guth, James. 1985. "The Christian Right Revisited: Partisan Realignment among Southern Baptist Ministers." Paper presented to the annual meeting of the Midwest Political Science Association, Chicago.

Keith, Bruce E., David B. Magleby, Candice J. Nelson, Elizabeth Orr, Mark Westlye, and Raymond E. Wolfinger. 1977. "The Myth of the Independent Voter." Paper presented at the annual meeting of the American Political Science Association, Washington, DC.

Key, V. O., Jr. 1949. *Southern Politics in State and Nation*. New York: Knopf.

Norpoth, Helmut. 1987. "Under Way and Here to Stay: Party Realignment in the 1980s?" *Public Opinion Quarterly* 51:376–91.

Patel, Kant, Denny Pilant, and Gary Rose. 1982. "Born-again Christians in the Bible Belt." *American Politics Quarterly* 10:255–72.

Swansbrough, Robert H., and David M. Brodsky, eds. 1988. *The South's New Politics: Realignment and Dealignment*. Columbia: University of Southern Carolina Press.

BRINGING STRONG TIES BACK IN
INTERPERSONAL GATEWAYS TO POLITICAL INFORMATION AND INFLUENCE

BRUCE C. STRAITS

Abstract With whom and how frequently do people discuss political matters? A theoretical model is formulated and tested using information on respondents' core social networks from the 1987 General Social Survey. The frequency of political discussions depends upon factors of demand (personal attributes) and supply (opportunities within social contexts). Educational attainment and age positively affect political discussion frequency, although indirectly through intervening political orientation and social contexts variables. Demand is positively related to political interest, strong party attachment, liberal views, and newspaper readership. The findings underscore the importance of including close relatives and friends (strong ties) in studying social contexts of political behavior. The opportunity structure (supply of potential discussants) is enhanced by the presence of a spouse and other adult household members, by being in the labor force, and by having close personal relationships with spatially proximate core associates.

The classic Columbia voting studies (Berelson, Lazarsfeld, and McPhee 1954; Lazarsfeld, Berelson, and Gaudet 1948) revealed the role of primary-group contacts (family, co-workers, and friends) as providers of political information and influence. Despite the importance of personal networks, little is known about the factors that promote and restrict the flow of political information and influence (Sheingold 1973, pp. 714–15; Weatherford 1982, pp. 119–20). How frequently do people discuss politics and with whom? This paper focuses on political discussions that occur between individuals and core members (major discussion partners) of their social milieus.

Previous studies of political discussion partners largely were designed to tap specific segments of personal networks, such as parents

BRUCE C. STRAITS is associate professor of sociology at the University of California, Santa Barbara. The author thanks Michael Dolny, M. Kent Jennings, and M. Stephen Weatherford for comments on an earlier draft. The data used in this paper were made available by the Inter-University Consortium for Political and Social Research.

Public Opinion Quarterly Volume 55:432–448 © 1991 by the American Association for Public Opinion Research
All rights reserved 0033-362X/91/5503-0005/$2.50

(Jennings 1983), spouses (Niemi, Hedges, and Jennings 1977), male friends (Laumann 1973), neighbors (Weatherford 1982), and political discussion partners who were not relatives (Huckfeldt and Sprague 1987, 1988). This study complements the prior research by extending the focus to include both kin and other core discussion partners and by examining the frequency of political discussions as well as the choice of discussants.

Conceptual Framework for Modeling Political Intercourse

The extent to which *ego* (the focal respondent) engages in political discussions depends upon: (1) ego's motivation (demand) for political intercourse as well as (2) the opportunity structure (supply) of potential discussion partners (*alters*) in ego's social milieus.

EGO-DEMAND FOR POLITICAL DISCUSSIONS

A person may initiate or maintain a political conversation to meet diverse needs or desires, including seeking or providing information and interpretation, exchanging attitudes and values, and conforming to or influencing others. Although diverse, these reasons are strongly dependent upon ego's political nature. Specifically, ego-demand will be modeled as a function of background characteristics (education, age, sex, race), political orientation (interest, partisanship, ideological self-placement), and media exposure (newspaper readership).[1]

Background characteristics.—Educational attainment is expected to be positively related to political discussion frequency because schooling imparts the knowledge, motivation, and cognitive skills necessary to understand and appreciate current political matters (Converse 1988, pp. 288–89; Wolfinger and Rosenstone 1980, p. 102) and promotes a lifelong openness to learning and information seeking (Hyman, Wright, and Reed 1975).

Wolfinger and Rosenstone (1980, pp. 58–60) argue that relevant political skills, motivation, and knowledge also may be acquired outside of school through accumulated life experiences, which may explain why voter turnout generally increases with age even though older Americans have less formal education. Although relevant experiences may intensify political interest over the life course, other age-related processes may reduce or change the pattern of political activities among older persons (Bennett and Bennett 1986, pp. 188–89; Jennings and Markus 1988). Consequently, to allow for a possible curvilinear relationship, political intercourse will be modeled as a polynomial function of age.

---

[1] Table 1 provides detailed descriptions of the variables.

The likely effects of the other background attributes, sex and race, are less clear. Political participation differences between men and women and between blacks and whites lately have become slight or negligible when background controls are applied (Bennett and Bennett 1986, pp. 186–91; Kinder and Sears 1985, p. 699; Uhlaner 1989, pp. 65–70).

Political orientation and media exposure.—Some of the impact of education and age-related experiences should be indirect, being mediated through political orientation (interest, partisanship, ideological self-placement) and media exposure, which are more proximate determinants of political intercourse. For example, schooling and life experiences should promote general interest in politics and national affairs, which in turn should heighten attentiveness to the political content of interpersonal and media channels.

Strength of partisanship (strong party attachment) and ideological self-placement (liberal-conservative spectrum) also are expected to intervene between background characteristics and demand for political intercourse. Previous studies repeatedly have found that strength of party attachment is strongly related to various modes of political participation (Bennett and Bennett 1986, pp. 191–92; Campbell et al. 1986). The literature on the relationship between ideological position and political participation is less clear-cut. Conservatives were more active during the 1956, 1960, and 1964 presidential elections, but the ideological pattern then reversed with greater participation of liberals in the 1972 and 1976 elections (Beck and Jennings 1979; Verba and Nie 1972).

Finally, newspaper readership should be positively related to political discussion frequency (Burstein 1972). Diminishing reliance on newspapers for campaign information appears to be one of the causes of declining voter turnout (Shaffer 1981; Teixeira 1987). Those who follow politics in newspapers should be motivated to share, exchange, and compare this information with others.

SUPPLY-SIDE OPPORTUNITIES FOR POLITICAL DISCUSSIONS

Ego's desire to participate in, or to avoid, political discussions is abetted or impeded by the supply of appropriate discussants in the surrounding social environment (Huckfeldt and Sprague 1987, 1988; Weatherford 1982, p. 131). A scarcity of suitable discussion partners will restrict the political intercourse of politically motivated citizens. Some less motivated individuals, on the other hand, may find themselves discussing politics more often than they desire to accommodate politically involved associates.

The opportunity structure for political discourse will be examined here both in terms of broad indicators of social context (marital status, number of adult household members, employment status, core network size, and geographical region) and specific ties between discussion partners and ego (role relationship, closeness, political party similarity).

Social contexts.—At home, the presence of a spouse and other adult household members should enhance the opportunities for political discourse. The observed similarity of husbands' and wives' political attitudes and actions such as turning out to vote (Straits 1990) results to some extent from such political exchanges (Niemi, Hedges, and Jennings 1977, pp. 142–44).[2] Even a spouse uninterested in politics will often have to discuss the subject with a politically motivated partner.

Outside the home, an individual may find suitable discussion partners among co-workers and other associates. Political discussion frequency is expected to be associated positively with working outside the home (Burstein 1972) and with core network size. Larger core networks tend to be more heterogeneous, thus linking ego to diverse contacts and information sources (Campbell, Marsden, and Hurlbert 1986). Finally, geographical region is included in the model to pick up possible regional variation in political discourse.

Discussion partner relationships.—Political intercourse between ego and a core associate is expected to occur more often when the relationship is close, when contact is frequent (spatial proximity), and when they share similar social characteristics and beliefs.

Compared to casual acquaintances, political discussions with close friends should be more beneficial and less costly (Huckfeldt and Sprague 1987; Weatherford 1982). Since ego is familiar with the political nature of close associates, ego can maximize the likelihood of achieving the desired conversational outcome by matching a political topic with an appropriate discussant. Political discourse with casual friends, on the other hand, may be less rewarding if ego misjudges the friends' political knowledge and leanings. Finally, close bonds are less likely to be undermined than more casual relationships by the disclosure of political ignorance, extreme views, and other potentially sensitive or embarrassing information (Hays 1989).

The frequency of political discourse between ego and an associate is expected to be related directly to the amount of time the two are in direct contact. When ego's memory, the media, or another person triggers a political theme, those close at hand provide convenient discussion partners. Finally, social as well as spatial proximity should promote political discourse. The greater ego's status similarity and political compatibility with an associate, the greater their frequency of political discussions (Huckfeldt and Sprague 1988; Rogers and Bhowmik 1971; Verbrugge 1977). Disclosure of potentially sensitive or embarrassing information should be less threatening to political discussants if they are of equal rather than disparate statuses.

---

[2] Spousal similarities also may reflect self-selection of like-minded mates and shared family circumstances (Niemi, Hedges, Jennings 1977, pp. 144–45).

Data, Measures, and Methods

The network data for this analysis are from the 1987 General Social Survey (GSS) in which respondents were asked: "From time to time, most people discuss *important matters* with other people. Who are the *people* with whom you discuss such matters?" The criterion choice of *important* matters is intended both to avoid undersampling certain relationships (e.g., asking about close friends undersamples kinship ties) and to elicit the core associates who influence the respondents' attitudes and behavior (Burt 1984; Marsden 1987). Respondents then were asked information about each of the first three named discussion partners, including how often political matters are discussed. Table 1 provides detailed descriptions of the dependent and independent variables.

Excluded from the analysis are about 5 percent of the respondents who either named no discussion partners or did not answer the question about political discussions with the first named partner (alter 1).[3] Separate dependent variables were constructed (*a*) to describe the reported frequency of political discussions with each named partner (alters 1–3) and (*b*) to summarize the total volume of political intercourse with these major discussion partners.

Scaling political discussion frequency with partners (alters 1–3) is difficult because of the wording of the response categories: "almost daily," "at least weekly," "at least monthly," "at least yearly," "less than yearly," and "never." For a response of "at least monthly," for example, one might assign the midpoint (31.5) of the implied range (12–51) of yearly contacts. On the other hand, response category choice may be strongly influenced by the salient minimum values ("daily," "weekly," "monthly," etc.), which suggests using these values to score the responses (e.g., 12 for "at least monthly"). The

---

[3] In response to the "discussed important matters" name generator, 3.8 percent named no one, 13.0 percent cited one discussion partner, 18.7 percent cited two, 50.5 percent cited three, 7.2 percent cited four, 3.4 percent cited five, and 3.4 percent cited six or more partners.

Table 1. Descriptions of Variables Used in the Analysis

Variable	Description
Political intercourse:	
Alter$_i$ discussion frequency	Coded 310 if reported discussing political matters with partner *i* "almost daily," 111.75 if "at least weekly," 21.75 if "at least monthly," 3.5 if "at least yearly," 0.26 if "less than yearly," 0 if "never" or "don't know," and missing if "no answer" or if fewer than *i* partners named.
Discussion index	Sum of above alter$_i$ discussion frequency (missing recoded as 0 for alters 2 and 3) for first three named discussion partners. Coded as missing if respondent named no discussion partners or didn't answer the political discussion question for alter 1.
Background characteristics:	
Age, age-squared	Coded in years and specified as a second-degree polynomial.
Education	Years of schooling completed.
Female	1 if female, 0 if male.
Racial self-identification:	
Black	1 if black, 0 if white or other.
Other	1 if other, 0 if white or black.
Political orientation and media exposure:	
Interest in politics:[a]	
Very interested	1 if "very interested," 0 if not.
Somewhat interested	1 if "somewhat interested," 0 if not.
Partisanship	Folded GSS variable PARTYID coded 4 if "strong" Republican or Democrat or "other party"; 3 if "not very strong" Republican or Democrat; 2 if "independent close to" Republican or Democrat; and 1 if "independent" or "don't know."

Table 1. (Continued)

Variable	Description
Ideological self-placement	GSS variable POLVIEWS: 7-point scale ranging from "extremely liberal" (coded 1) to "extremely conservative" (coded 7). "Don't know" responses were coded 4.
Newspaper readership	1 if read a newspaper "every day," 0 if less frequently.
Social contexts:	
Married	1 if currently married, 0 if not.
Adults in household	Number of adults (over 17 years old) living in the household.
Employed	Based on GSS variable WRKSTAT: 1 if currently in the labor force (employed or unemployed); 0 if retired, in school, keeping house, or other work status.
Core network size	GSS variable NUMGIVEN: number of discussion partners named. The "6+" category has been recoded as 6.5 (Marsden 1987, p. 125, n. 8).
Region:[b]	
Pacific	1 if Pacific division, 0 if not.
South and West	1 if South Atlantic, East South Central, West South Central, West North Central, or Mountain division, 0 if not.
Discussion partner relationships:	
Closeness$_i$	Scaling suggested by Burt and Guilarte (1985, pp. 5–6): coded 1 if ego feels "equally close" to all named discussion partners (GSS variable EQCLOSE) or if ego feels "especially close" to alter *i* (GSS variable RCLOSE$_i$); otherwise coded 0.7.
Party dissimilarity$_i$	1 if reported party identification (Democrat, Republican, independent, other) for ego and alter *i* differ, 0 if the same (GSS variables PARTYID and PARTYID$_i$).

Role relationships:[c]	
Spouse$_i$	1 if alter *i* is a spouse or "person living with as if married," 0 if not.
Parent$_i$	1 if alter *i* is a parent, 0 if not.
Sibling$_i$	1 if alter *i* is a brother or sister, 0 if not.
Child$_i$	1 if alter *i* is a child, 0 if not.
Other family$_i$	1 if alter *i* is some other family member (e.g., "grandparent, grandchildren, cousins, aunts, uncles, nephews, nieces, in-laws"), 0 if not.
Co-worker$_i$	1 if alter *i* is a co-worker ("work with or usually meet while working"), 0 if not.
Group member$_i$	1 if alter *i* is a member of a group to which ego belongs (e.g., "attends your church, or whose children attend the same school as your children, or belongs to the same club, classmate"), 0 if not.
Neighbor$_i$	1 if alter *i* is a neighbor, 0 if not.
Friend$_i$	1 if alter *i* is someone with whom ego gets "together for informal social occasions" (including "boyfriends" and "girlfriends"), 0 if not.
Adviser$_i$	1 if alter *i* is "a trained expert" ego turns "to for advice" (e.g., "a lawyer or clergyman"), 0 if not.
Other$_i$	1 if alter *i* has some other connection to the respondent, 0 if not.

NOTE.—See NORC (1987) for additional details on the 1987 GSS variables.

[a] Based on GSS variable INTPOL: "How interested are you in politics and national affairs?" Omitted categories: "only slightly interested," "not at all interested," "don't know."

[b] Omitted U.S. Bureau of Census divisions: New England, Middle Atlantic, and East North Central.

[c] Question eliciting above role relationships: "Here is a list of some of the ways in which people are connected to each other. Some people can be connected to you in more than one way. For example, a man could be your brother and he may belong to your church and be your lawyer. When I read you a name, please tell me *all* the ways that person is connected to you."

solution adopted here to measure alter discussion frequency involves giving equal weight to the midpoint and the minimum value of the implied range by averaging the two values (e.g., 21.75 for "at least monthly").[4]

The summary discussion index is operationalized by adding the above-scored discussion frequencies for the first three named partners. If the respondent's core network is of size 1 or 2, the discussion index is based on the 1 or 2 named alters. Previous use of this core network generator in the 1985 GSS revealed that discussion partners were cited in descending order of importance to the respondent (Burt 1986). Consequently, the discussion index measures the volume of political exchanges that occur between the respondents and their closest discussion partners.[5]

Ordinary least squares (OLS) regression, weighted to adjust for the 1987 GSS black oversample, is used to model political intercourse. Although it is likely that political exchanges with an alter will influence discourse with the other discussion partners, isolating the alters in separate models vastly simplifies estimation and interpretation. Similarly, the modeling assumes asymmetrical causal relationships among the other included variables; for example, that partisanship or newspaper readership influences political intercourse, but not vice versa. Again, reality is simplified to permit hypothesis testing with the limited data at hand.

Results

DISCUSSIONS WITH CORE ASSOCIATES

Table 2 reports results for two OLS regression models predicting the volume of political conversations with the respondent's core discussion partners. Model 1, which is limited to background variables, supports the hypothesized positive effects of age and education on political interactions. The metric coefficient for education indicates that the discussion index increases by slightly over eight units with each additional year of schooling. As hypothesized, political intercourse is a parabolic (quadratic) function of age. According to estimated model 1, the frequency of political conversations rises with age to a peak around age 59, and then decreases slightly with advancing age.[6] With cross-sectional data, however, it is not possible to determine the extent to which this curvilinear age pattern reflects life-course changes or cohort (generational) differences in political involvement.

---

[4] The responses "almost daily" and "less than yearly" were assigned minimum values of 292 (80 percent × 365) and 0.01, respectively.

[5] Since the index is a function of core network size, the latter is included in the analysis as an independent variable.

[6] The age inflection point was calculated by setting the first derivative equal to zero (Stolzenberg and Land 1983, p. 641).

Although the effects of sex and race are statistically insignificant in model 1, the coefficient for blacks becomes larger and significant when the political orientation and social contexts variables are included as shown in estimated model 2. Education and age, conversely, do not directly affect the volume of political conversations after the intervening political orientation and social contexts variables are included in the model.

The results confirm the hypothesized effects of political orientation and media exposure. Predicted discussion frequency increases by 71 units if the respondent is "very interested" in politics, by 58 units if party attachment is "strong," and by 30 units if a newspaper is read "every day." The effect of ideology, for which no hypothesis was offered, was also significant: liberals were more likely than conservatives to engage in political conversations.

Opportunity structure hypotheses were also supported. The estimated discussion index increases approximately by 25 units for each discussion partner named (core network size), by 14 units for each adult in the household, and by 27 units if the respondent is currently in the labor force. To test the hypothesis that political intercourse should flourish under conditions of high demand and discussant availability, interaction terms were formed from the product of the political interest and marital status dummy variables.[7] The results confirm the prediction: the combination of being "very interested" in politics and currently married translates into a 65-unit increase in estimated political discussion frequency.

Finally, the frequency of political conversations varied by geographical region. Compared to respondents living in New England, Middle Atlantic, or East North Central, politics were discussed more frequently by residents of the Pacific division and less frequently by residents living elsewhere in the United States.[8]

DISCUSSION PARTNER RELATIONSHIPS

Discussion partner differences in political contact are summarized in table 3. The background, political orientation, and media exposure

---

[7] For example, "Married × Very Interested" is coded 1 if the respondent is currently married and very interested in politics. Otherwise, it is coded 0.

[8] The census divisions were represented by a set of dummy variables in preliminary analyses. Regions that were not significantly different from each other (F-test) were then combined as shown in table 1. The regional effects may reflect differences in social environment not otherwise represented by variables in the model.

Table 2. OLS Regression Analysis of Political Discussion Index

Variable	Model 1 (*N* = 1,712) Coefficient	Model 1 (*N* = 1,712) *t*-value	Model 2 (*N* = 1,684) Coefficient	Model 2 (*N* = 1,684) *t*-value	Descriptive Statistics Mean	Descriptive Statistics S.D.
Discussion index	. . .	. . .	. . .	. . .	142.74	198.43
Background characteristics:						
Age	4.568	3.01**	1.392	.91	44.92	17.53
Age-squared	−.039	−2.58**	−.012	−.78	2,324.65	1,775.95
Education	8.270	5.03***	−.612	−.36	12.60	3.15
Female[a]	−18.076	−1.88	4.113	.43	.57	.50
Racial self-identification:						
Black[a]	16.226	1.10	35.094	2.41*	.12	.32
Other[a]	23.112	.89	23.643	.95	.04	.18
Political orientation and media exposure:						
Interest in politics:						
Very interested	. . .	. . .	71.283	3.82***	.23	.42
Somewhat interested	. . .	. . .	31.856	2.03*	.43	.50

Partisanship	. . .	. . .	14.557	2.92**	2.87	.96
Ideological self-placement[a]	. . .	. . .	−9.040	−2.58*	4.04	1.30
Newspaper readership	. . .	. . .	30.273	3.08**	.55	.50
Social contexts:						
Married	. . .	. . .	3.372	.21	.55	.50
Married × very interested	. . .	. . .	65.178	2.64**	.12	.33
Married × somewhat interested	. . .	. . .	23.435	1.12	.25	.43
Adults in household	. . .	. . .	13.650	2.26*	1.98	.81
Employed	. . .	. . .	27.218	2.29*	.68	.47
Core network size	. . .	. . .	25.350	5.98***	2.75	1.11
Region:						
Pacific[a]	. . .	. . .	60.586	4.05***	.13	.33
South and West[a]	. . .	. . .	−39.665	−4.04***	.49	.50
Intercept	−68.881	−1.68	−62.688	−1.31		
R^2 (adjusted)	.02		.14			

NOTE.—Results weighted by GSS variable OVERSAMP (NORC 1987, p. 441).

[a] Two-tailed tests. Otherwise, p values are for one-tailed tests.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 3. OLS Regression Analysis of Discussions with Each Partner

	Alter 1 (*N* = 1,684)		Alter 2 (*N* = 1,432)		Alter 3 (*N* = 1,097)	
Variable	Coefficient	*t*-value	Coefficient	*t*-value	Coefficient	*t*-value
Background characteristics:						
Age	.146	.18	.268	.40	.961	1.20
Age-squared	−.000	−.01	−.003	−.44	−.014	−1.72*
Education	1.097	1.24	−.939	−1.22	−1.100	−1.22
Female[a]	5.802	1.18	−3.073	−.73	1.181	.25
Racial self-identification:						
Black[a]	18.684	2.41*	16.362	2.43*	20.753	2.57*
Other[a]	−2.524	−.19	2.816	.27	30.213	2.47*
Political orientation and media exposure:						
Interest in politics:						
Very interested	28.776	3.30***	35.584	5.79***	25.002	3.55***
Somewhat interested	12.351	1.76*	17.146	3.41***	9.452	1.63
Partisanship	4.628	1.76*	5.105	2.26*	9.816	3.87***
Ideological self-placement[a]	−4.727	−2.57*	−3.190	−2.05*	−2.687	−1.49
Newspaper readership	12.661	2.46**	7.566	1.73*	16.228	3.23***
Social contexts:						
Core network size	4.769	2.12*	2.914	1.28	−1.346	−.46
Region:						
Pacific[a]	32.843	4.18***	20.525	3.17**	7.419	1.04
South and West[a]	−14.866	−2.86**	−16.821	−3.83***	−14.951	−2.96**

Discussion partner relationships:						
Closeness	97.188	3.24***	40.289	2.62**	40.313	2.38**
Party dissimilarity	−1.538	−.28	13.806	3.02**	2.691	.52
Role relationships:						
Spouse	20.561	2.07*	50.054	3.37***	29.773	1.42
Spouse × very interested	66.938	5.02***	68.499	2.27*	39.558	1.09
Spouse × somewhat interested	27.783	2.44**	−4.727	−.23	45.533	1.76*
Parent[a]	−2.208	−.24	−9.085	−1.41	−4.292	−.54
Sibling[a]	−17.254	−1.57	1.045	.13	6.381	.73
Child[a]	−4.282	−.40	−9.531	−1.34	9.304	1.05
Other family[a]	−5.446	−.47	1.731	.22	7.927	.91
Co-worker	26.358	3.78***	22.703	4.01***	40.183	6.35***
Group member[a]	−5.475	−.80	−14.995	−2.59*	−16.029	−2.37*
Neighbor[a]	15.085	1.70	4.401	.71	10.589	1.51
Friend[a]	−.843	−.14	.021	.00	6.128	1.01
Adviser[a]	−16.786	−3.11**	−.838	−.18	−1.802	−.34
Other[a]	15.753	1.10	4.388	.26	29.814	1.91
Intercept	−83.573	−2.12	−15.861	−.60	−35.997	−1.24
R^2 (adjusted)	.17		.11		.13	

NOTE.—Results weighted by GSS variable OVERSAMP (NORC 1987, p. 441).

[a] Two-tailed tests. Otherwise, p values are for one-tailed tests.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

variables generally have the same effects across the discussion partners.[9] An exception is partisanship, the impact of which increases from the first to the third partner. Presumably, strong party identifiers are more likely than weak identifiers and independents to discuss politics with less spatially and socially proximate associates.[10]

Core network size was included in the alter models to explore a contagion hypothesis that a larger and presumably more heterogeneous network increases the flow of political information to ego, which in turn triggers additional interpersonal transmissions. The hypothesis is weakly supported for the first discussion partner.[11] Also, the reported frequency of political conversations with each partner are moderately correlated (product-moment correlations between alter *i* and *j* range from .48 to .53).

As hypothesized, the respondents discuss politics more frequently with closer associates. The metric coefficient for alter 1, for example, indicates that discussion frequency decreases by about 29 units (0.3 × 97.2) if ego felt "less close" to this partner than to others. The hypothesis that political conversations will occur more often with politically compatible discussion partners was not supported when incompatibility was measured by ego-alter political party dissimilarity.[12]

The effects of discussants' role relationships to ego lend support to the social and spatial proximity hypotheses. The respondents engaged more frequently in political conversations with spatially proximate associates (spouses and co-workers).[13] The results also suggest that political intercourse is less likely to occur if there are social or status differences between ego and alters, such as discussion partners who are advisers, parents, or children. Also, political conversations are less frequent if the partner relationship is based on joint membership in an organization or interest group.

The political interest and spouse interaction terms highlight the importance of the opportunity structure for political discourse. The combination of being "very interested" in politics and naming a spouse as alter 1 translates into a 67-unit increase in estimated political discussion frequency.[14] The results also indicate that individuals who are only slightly or not interested in politics are significantly more likely to engage in political conversations if a spouse is named as a discussion partner.

---

[9] Some of the previously used broad measures of social context—marital status, number of adult household members, and employment status—have been excluded since more direct information is available on the discussant's role relationship to ego (spouse, co-worker, etc.).

[10] Discussion partner order is consistent with the previous use of the network questions in the 1985 GSS (Burt 1986) in that spouses tended to be named first (39 percent of alter 1 choices; 4 percent of alters 2 and 3) and discussion partners were named in descending order of closeness and political discussion frequency.

[11] This test is flawed since network size increases and variability decreases with discussion partner order.

[12] The estimated coefficient is in the opposite direction than hypothesized for alter 2, and nonsignificant for the others.

[13] The 1985 GSS respondents reported more contact with spouses and co-workers than with other discussion partners.

Discussion

By focusing on the frequency of political conversations within core networks, this research provides a different perspective that complements previous studies of political discussion partners. Unlike the neighborhood-based networks studied by Weatherford (1982, pp. 129–30), within core networks, closeness of personal relationships and spatial proximity (frequency of contact) are important determinants of political intercourse. Unlike Huckfeldt and Sprague's (1987, 1988) finding that people choose politically compatible discussion partners, political congeniality appears to be unimportant when the network includes relatives and other close associates and when discussion frequency rather than choice of discussant is modeled.

The findings underscore the importance of including close relatives and other strong ties in studying social contexts of political behavior. Even though spouses may be quite similar in education, age, and other social characteristics, their network gateways through which political information and influence flow may be quite different. The results suggest, for example, that a married-couple's exposure to political information is enlarged by co-worker links to workplace environments. Also, individuals uninterested in politics nevertheless may receive considerable exposure to political information from politically motivated close associates.

References

Beck, Paul Allen, and M. Kent Jennings. 1979. "Political Periods and Political Participation." *American Political Science Review* 73:737–50.

Bennett, Stephen E., and Linda L. M. Bennett. 1986."Political Participation." In *Annual Review of Political Science*, ed. Samuel Long, 1:157–204. Norwood, NJ: Ablex.

Berelson, Bernard, Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting*. Chicago: University of Chicago Press.

Burstein, Paul. 1972. "Social Structure and Individual Political Participation in Five Countries." *American Journal of Sociology* 77:1087–1110.

Burt, Ronald S. 1984. "Network Items and the General Social Survey." *Social Networks* 6:293–339.

———. 1986. "A Note on Sociometric Order in the General Social Survey Network Data." *Social Networks* 8:149–74.

---

[14] The total impact (main effects and interaction) of alter 1 spouse and "very interested" is 116.3 discussion units (20.6 + 28.8 + 66.9).

Burt, Ronald S., and Miguel G. Guilarte. 1985. *A Note on Scaling the General Social Survey Network Item Response Categories*. GSS Technical Report no. 61. Chicago: NORC.

Campbell, Karen E., Peter V. Marsden, and Jeanne S. Hurlbert. 1986. "Social Resources and Socioeconomic Status." *Social Networks* 8:97–117.

Campbell, James E., Mary Munro, John R. Alford, and Bruce A. Campbell. 1986. "Partisanship and Voting." *Research in Micropolitics* 1:99–126.

Converse, Philip E. 1988. "Perspectives on the Democratic Process." *Michigan Quarterly Review* 27:285–99.

Hays, Robert B. 1989. "The Day-to-Day Functioning of Close versus Casual Friendships." *Journal of Social and Personal Relationships* 6:21–37.

Huckfeldt, Robert, and John Sprague. 1987. "Networks in Context: The Social Flow of Political Information." *American Political Science Review* 81:1197–1216.

———. 1988. "Choice, Social Structure, and Political Information: The Informational Coercion of Minorities." *American Journal of Political Science* 32:467–82.

Hyman, Herbert H., Charles R. Wright, and John S. Reed. 1975. *The Enduring Effects of Education*. Chicago: University of Chicago Press.

Jennings, M. Kent. 1983. "Gender Roles and Inequalities in Political Participation: Results from an Eight-Nation Study." *Western Political Quarterly* 36:364–85.

Jennings, M. Kent, and Gregory B. Markus. 1988. "Political Involvement in the Later Years: A Longitudinal Survey." *American Journal of Political Science* 32:302–16.

Kinder, Donald R., and David O. Sears. 1985. "Public Opinion and Political Action." In *The Handbook of Social Psychology*, ed. Gardner Lindzey and Elliot Aronson, 3d ed., 2:659–741. New York: Random House.

Laumann, Edward O. 1973. *Bond's of Pluralism: The Form and Substance of Urban Social Networks*. New York: Wiley.

Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet. 1948. *The People's Choice*, 2d ed. New York: Columbia University Press.

Marsden, Peter V. 1987. "Core Discussion Networks of Americans." *American Sociological Review* 52:122–31.

Niemi, Richard G., Roman Hedges, and M. Kent Jennings. 1977. "The Similarity of Husbands' and Wives' Political Views." *American Politics Quarterly* 5:133–48.

NORC. 1987. *General Social Surveys, 1972–1987: Cumulative Codebook*. Chicago: NORC.

Rogers, Everett M., and Dilip K. Bhowmik. 1971. "Homophily-heterophily: Relational Concepts for Communication Research." *Public Opinion Quarterly* 34:523–38.

Shaffer, Stephen D. 1981. "A Multivariate Explanation of Decreasing Turnout in Presidential Elections, 1960–1976." *American Journal of Political Science* 25:68–95.

Sheingold, Carl A. 1973. "Social Networks and Voting: The Resurrection of a Research Agenda." *American Sociological Review* 38:712–20.

Stolzenberg, Ross M., and Kenneth C. Land. 1983. "Causal Modeling and Survey Research." In *Handbook of Survey Research*, ed. Peter H. Rossi, James D. Wright, and Andy B. Anderson. New York: Academic Press.

Straits, Bruce C. 1990. "The Social Context of Voter Turnout." *Public Opinion Quarterly* 54:64–73.

Teixeira, Ruy A. 1987. *Why Americans Don't Vote: Turnout Decline in the United States, 1960–1984*. New York: Greenwood.

Uhlaner, Carole J. 1989. "Turnout in Recent American Presidential Elections." *Political Behavior* 11:57–79.

Verba, Sidney, and Norman H. Nie. 1972. *Participation in America: Political Democracy and Social Equality*. New York: Harper & Row.

Verbrugge, Lois M. 1977. "The Structure of Adult Friendship Choices." *Social Forces* 56:576–97.

Weatherford, M. Stephen. 1982. "Interpersonal Networks and Political Behavior." *American Journal of Political Science* 26:117–43.

Wolfinger, Raymond E., and Steven J. Rosenstone. 1980. *Who Votes?* New Haven, CT: Yale University Press.